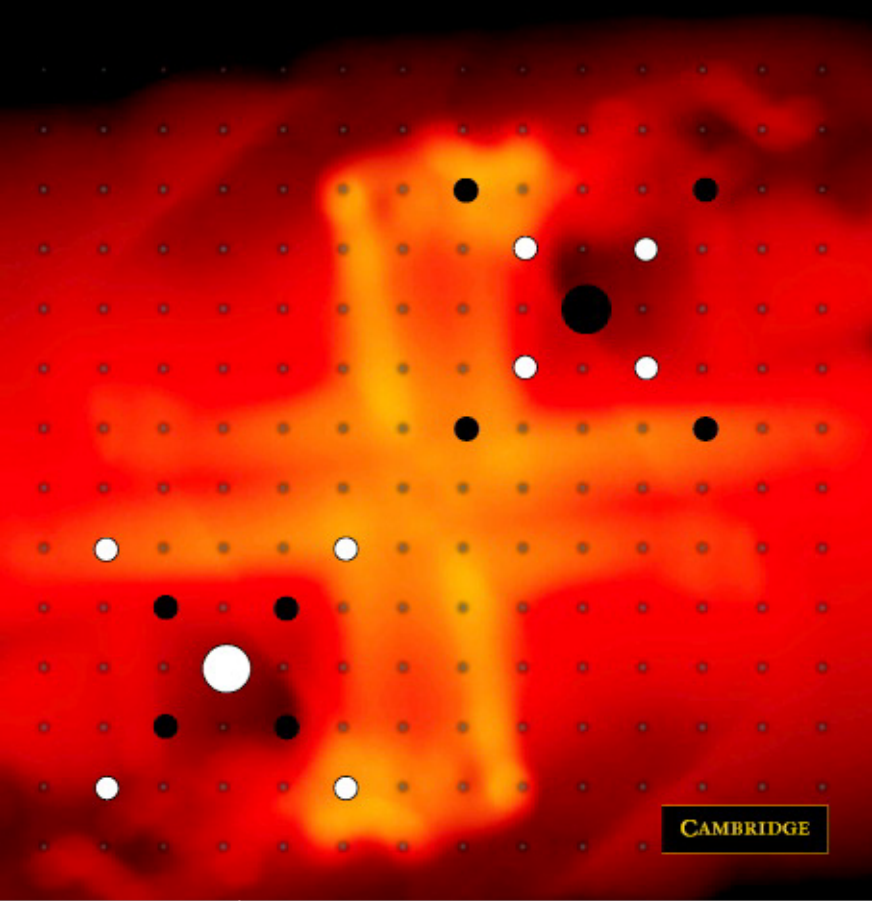


Optical Processes in Solids

Yutaka Toyozawa



CAMBRIDGE

CAMBRIDGE

more information - www.cambridge.org/9780521554473

This page intentionally left blank

OPTICAL PROCESSES IN SOLIDS

It is known from theory and from comparison with experiment that optical spectra and their lineshapes can tell us much about the microscopic motion of the constituent particles of matter – electrons and nuclei.

A unifying element that links the apparently diverse phenomena observed in optical processes is the dielectric dispersion of matter. It describes the response of matter to incoming electromagnetic waves and charged particles, and thus predicts their behavior in the self-induced field of matter, known as polariton and polaron effects. The energies of phonon, exciton and plasmon, quanta of the collective motions of charged particles constituting the matter, are also governed by the dielectric dispersion. Since the latter is a functional of the former, one can derive useful relations for their self-consistency. The nonlinear response to laser light inclusive of multiphoton processes, and excitation of atomic inner shells by synchrotron radiation, are also described in the same context.

Within the validity of the configuration coordinate model, photo-induced lattice relaxation and chemical reaction are described in terms of the interaction modes. Applying these arguments equally to both ground and relaxed excited states, one obtains a novel and global perspective on structural phase transitions as well as on the nature of interatomic bonds.

The book will be of value to graduate students and researchers in solid state physics, physical chemistry, materials science and biochemistry.

YUTAKA TOYOZAWA is a Professor Emeritus at the University of Tokyo. He received B.Sc. and D.Sc. degrees from the University of Tokyo. He worked 5 years at the Research Institute for Fundamental Physics (Yukawa Hall) of Kyoto University, for 27 years at the Institute for Solid State Physics of the University of Tokyo, with the last 3 years as Director, and for 10 years at the Physics Department of Chuo University. He also spent some time at Technische Hochschule Stuttgart (1962–3), the University of Illinois (1968) and the University of Rochester (1968) as a visiting professor. Professor Toyozawa has made a number of pioneering contributions to the theoretical physics of light–matter interaction in condensed matter, and is best known for his work on the exciton. His own philosophy, inclusive of novel ideas and concepts, has had a profound influence on young scientists, experimental as well as theoretical, through his presentations at international schools and meetings, numerous original review papers, and books such as *Physics of Elementary Excitations* (with S. Nakajima and R. Abe) and *Excitonic Processes in Solids* (with M. Ueta, H. Kanzaki, K. Kobayashi, and E. Hanamura). He has received a number of awards for his scientific contributions, such as the Nishina memorial prize (1966) and the medal with purple ribbon (1989). Since his retirement in 1997, he is still working actively in developing the interdisciplinary areas.

OPTICAL PROCESSES IN SOLIDS

YUTAKA TOYOZAWA



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press

The Edinburgh Building, Cambridge CB2 2RU, United Kingdom

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9780521554473

© Cambridge University Press 2003

This book is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2003

ISBN-13 978-0-511-07742-5 eBook (NetLibrary)

ISBN-10 0-511-07742-4 eBook (NetLibrary)

ISBN-13 978-0-521-55447-3 hardback

ISBN-10 0-521-55447-0 hardback

ISBN-13 978-0-521-55605-7 paperback

ISBN-10 0-521-55605-8 paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this book, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

<i>Preface</i>	<i>page</i> ix
<i>Acknowledgments</i>	xii
<i>List of principal symbols</i>	xiii
1 Radiation field	1
1.1 Maxwell equations for electromagnetic field	1
1.2 Electromagnetic wave	3
1.3 Canonical equations of motion for electromagnetic waves	4
1.4 Thermal radiation field	7
2 Quantum mechanics	10
2.1 Elements of quantum mechanics	10
2.2 Quantum mechanics of a harmonic oscillator	15
2.3 Uncertainty relation and complementarity	19
2.4 Density matrix	23
3 Interaction of radiation with matter	28
3.1 A and B coefficients of Einstein	28
3.2 Relation between absorption and emission spectra in thermal equilibrium	30
3.3 Canonical form of radiation–matter interaction	31
3.4 First-order optical processes: absorption and emission of photon	33
3.5 Quantal vs. classical field description of absorption process	35
3.6 Oscillator strength and the sum rule	37
4 Electronic vs. nuclear motions and the optical spectra of localized electrons in solids	40
4.1 Born–Oppenheimer approximation	40
4.2 Classical approximation of vibrational motion	42
4.3 Perturbation theory in electron–vibration interaction	43
4.4 Relaxation of nuclear positions after electronic transition	45
4.5 Vibrational structures in absorption spectra	48

4.6	Emission vs. absorption spectra	52
4.7	Interaction modes in degenerate excited states	53
4.8	Phonon coupling strength of localized electrons in solids	64
5	Lattice vibrations	69
5.1	Simple crystal lattices	69
5.2	Lattice vibrations in general crystals	73
5.3	Interaction of optical modes with electromagnetic field	77
5.4	Localized and quasi-localized modes	83
5.5	Quantum mechanics of lattice vibrations	86
6	Electric susceptibility and dielectric constant	89
6.1	Linear response theory for electric susceptibility	89
6.2	Analytic properties, dispersion relation and sum rule	95
6.3	Relation between electric susceptibility and dielectric constant	96
6.4	Dielectric constant in terms of phonons	99
6.5	Plasma oscillation and screening effect of electron gas	100
6.6	Energy loss of a charged particle injected into a dielectric	102
6.7	Optical activity	103
7	One-electron states in solids	107
7.1	Energy band	107
7.2	Band model for the classification of solids	109
7.3	The roles of impurity atoms	109
7.4	Effective mass approximation for shallow impurity states	110
7.5	Isoelectronic impurity	112
8	Excitons	113
8.1	Frenkel excitons	113
8.2	Wannier excitons	117
8.3	Elementary excitations in many-electron system of insulator	118
8.4	Optical spectra of excitons	123
8.5	Spin-orbit and exchange interactions in a crystal field	128
8.6	Polaritons and optical constants	132
8.7	Coherent potential approximation for exciton in mixed crystals	137
8.8	Giant oscillator strength of an impurity-bound exciton	141
8.9	Amalgamation- and persistence-type excitons in mixed crystals	143
9	Polaron and the self-trapped state	149
9.1	Polarons in ionic crystals	149
9.2	Polaron theories for intermediate to strong coupling ranges	155
9.3	Electronic polaron, a renormalized electron in a many-electron field	162
9.4	Short-range electron-phonon interaction and self-trapping	168
9.5	Continuum model for self-trapping	176

9.6	Microscopic structures of self-trapped electrons, holes and excitons	187
10	Optical spectra of exciton in the phonon field	192
	Overview	192
10.1	Exciton–phonon interaction and the polaron effect of an exciton	194
10.2	Phonon-induced interference in multicomponent exciton spectra	196
10.3	Motional narrowing of exciton linewidth	200
10.4	Motional reduction of phonon sidebands and the zero-phonon line	204
10.5	Phonon sidebands of vibronic excitons in molecular crystals	209
10.6	Correlation of absorption and emission spectra of excitons in the phonon field	213
10.7	Examination of model Hamiltonians	230
10.8	Polariton bottleneck for exciton–photon conversion	234
11	Higher-order optical processes	239
11.1	Nonlinear responses and multiphoton processes	239
11.2	Two-photon spectroscopy	242
11.3	Light scattering	247
11.4	Resonant light scattering by a localized electron	254
11.5	Correlated absorption and emission of a localized electron	262
11.6	Resonant light scattering in an exciton–phonon system	271
11.7	Giant two-photon absorption and hyper-Raman scattering associated with an excitonic molecule	279
11.8	Seeking materials with large nonlinear susceptibilities	282
12	Inner-shell excitation	284
12.1	Inner-shell electrons and synchrotron radiation	284
12.2	Hierarchical structures in absorption spectra of inner shells	286
12.3	Auger process and its competition with the radiative process	293
12.4	Lattice relaxation of core hole as reflected in recombination luminescence spectra	296
12.5	Photoelectron spectroscopy	298
12.6	Many-body effects in normal metals	304
12.7	Core-electron spectra of atoms with incomplete outer shells	308
13	Photo-induced structural changes	319
	Overview	319
13.1	Role of p-holes in formation of covalent bonds	322
13.2	Photo-chemical processes in alkali halides	326
13.3	Lattice relaxation of exciton across bandgap and its radiative and nonradiative annihilation	329

13.4	Neutral–ionic phase transition in quasi-one-dimensional charge transfer compounds	335
13.5	Photo-induced phase transition in TTF–CA	341
13.6	T – U – S model for N -electron N -site systems	345
13.7	Extrinsic self-trapping, and shallow–deep bistability of donor state	359
14	Light, matter and life (A scenario for the origin of life and its evolution)	371
14.1	Introduction	371
14.2	Energy source for creating living matter and sustaining living activities	372
14.3	Conversion of radiation energy to chemical energy	375
14.4	Kinetic equations for chemical evolution under solar radiation	378
14.5	Origin of natural selection and specific structure of biomolecules	382
14.6	The roles of water in photochemical reactions	383
14.7	Some remarks and further speculations	387
	<i>Appendix 1</i> Derivation of exciton Hamiltonian from many-body Hamiltonian	391
	<i>Appendix 2</i> Transient spectra for second-order optical processes	395
	<i>References (by chapter)</i>	398
	<i>Subject index</i>	415

Preface

Light is the most important medium of recognition in human life, as is obvious from the prominent role of vision among the five human senses. In its most naive form, one can recognize by light the position and shape of an object on the one hand and its color on the other. From the time that light was revealed to be a type of waveform, these two functions have been ascribed to the *spatially* and *temporally* oscillating factors of the wave, and have been applied to the diffraction study of the microscopic *structures* of matter and the spectroscopic study of microscopic *motions* in matter, respectively. It was Maxwell who revealed light to be an electromagnetic wave, which could be deduced from his fundamental equations for an electromagnetic field in a vacuum (Chapter 1).

Einstein, in his theory of relativity, incorporated three-dimensional space and one-dimensional time into a four-dimensional spacetime coordinate system on the basis of his Gedanken experiment with light as a standard signal for communication between observers. In contradistinction to the Newtonian equation of motion, the Maxwell equations survived this revolution of physics, proving themselves to be invariant against Lorentz transformation among different coordinate systems in relative motion. Thus, light acquired the position of universality governing the space-time framework itself, something more than an unknown and intangible medium which was once called “the ether”.

Another important role of light in the history of physics is the spectral distribution of a radiation field in thermal equilibrium with matter. Planck’s *working hypothesis* introduced ad hoc to remove the deviation of the prediction of classical mechanical theory from observation, together with the empirical constant h now called by his name, turned out to be a doorway to the discovery of quantum mechanics (Chapter 2). The first step beyond Planck was made by Einstein who explained the Franck–Hertz experiment on photoelectrons through his hypothesis of a *quantum of light* – a photon as it is now called – which was a *model substantiating* Planck’s working hypothesis but was in a sense also a revival of Newton’s corpuscular model

of light discarded long before. Compton explicitly associated *momentum* as well as *energy* to this hypothetical *particle* in conformity with the well-known energy–momentum relationship in the electromagnetic *wave*, and successfully explained his X-ray electron scattering experiment by applying the energy–momentum conservation law to two colliding particles. The success of two *apparently contradicting* models – wave and particle – depending upon the phenomena to be explained, was increasingly embarrassing to the physicists of that time.

What was found with light (wave \rightarrow particle) took place with an electron in the opposite direction (particle \rightarrow wave). Thus, de Broglie advocated that an electron with momentum p and energy E behaves like a wave with wave number $k = p/\hbar$ and angular frequency $\omega = E/\hbar$, the relation advocated by Einstein and Compton for a quantum of light. Shortly afterwards, Schrödinger presented his famous *wave equation* on the basis of this idea.

Another approach to quantum mechanics had been prepared by Bohr. He succeeded in explaining the series of lines in the spectrum of a hydrogen atom by assuming that the angular momentum of the orbital motion of an electron around a proton takes only integral multiples of \hbar and that the spectral lines correspond to the energy *difference* between these quantized stationary states. The *correspondence principle* he advocated, that the notion of the *transition* energy between (not the energies themselves of) the stationary states corresponds to the frequency of *classical* motion (as will most clearly be seen with a harmonic oscillator in Section 2.2), paved the way towards the discovery of *matrix mechanics* due to Heisenberg, which was soon proved to be equivalent to Schrödinger’s wave mechanics.

In this way, the interplay of radiation field and material particle, together with the wave–particle duality common to them, was the most important motivation driving classical physics towards the discovery of quantum mechanics. Atomic and molecular spectroscopy played the role of confirming quantum mechanics to the extremity of accuracy, particularly in the early days of its application.

Subsequent application of spectroscopic methods to condensed matter, which will be the main subject of the present book, opened an extremely fertile area of study which had never been expected, partly because light played the role of not only *probing* but sometimes *changing* the nature of the interatomic bonds of condensation – evolving from a *spectator* to a *participant*. First of all, light revealed the individual (Chapter 7) and collective (Chapters 5 and 8) motions of the constituent charged particles of matter which interact with various strengths and over different ranges. These motions are excited by electromagnetic waves, and the coefficient of the linear part of the response can be described in terms of the *dielectric constant* or conductivity *as a function of frequency* (Chapter 6). Even the *motion* of a charged particle and the *effective interaction* between charged particles within matter as the dielectric medium can be described in terms of a dielectric

function (Section 6.5; Chapters 8 and 9). The great difference in the rapidity of motion between electrons and nuclei allows us to apply the adiabatic approximation to describe their correlated motions (Chapter 4), providing the *configuration coordinate model with interaction mode* as a powerful tool to describe intuitively the rearrangement of nuclear positions after optical transition of (i) electrons in a molecule, or (ii) localized electrons in solids, inclusive of their photostructural *changes* and photochemical *reactions* (Chapter 13) – the light is here a participator in changing the nature of the interatomic bond. The way the dynamics of interacting collective excitations is reflected in typical optical spectra can be described systematically (Chapters 10, 11, and 12). Such studies turn out to be quite useful in judging what is going on in the microscopic world from a glance at the spectra – the light is here a spectator. The reader will find that the importance of the role of light as a participator increases in the order Chapter 11 through to Chapter 13.

On the basis of all that has been established in the main part of the book, the author presents in the last chapter (Chapter 14) a *tentative* scenario for the origin of life and its evolution, as a small contribution from physical science to a highly interdisciplinary problem yet unsolved. According to this scenario, solar radiation is a “creator” of life, far beyond being a mere participator.

The problems covered by this book are so diverse that presentation of some common viewpoints or intuitive models is desirable in order to describe them in a unified way and to correlate them in a plausible way. As such, we will take the dielectric description on the one hand and the configuration coordinate model with interaction mode (which can be extended to describe a phase transition) on the other (as already mentioned) – which are the two main features characterizing the book. At the same time, in a particular chapter we will refer not only to preceding but also to subsequent chapters in correlating the underlying physics and in showing instructive examples. In fact, such a style has already been taken in this preface. Also, we have tried to include many examples of comparison between theories and experiments which are considered to be of particular importance in the spectroscopic studies. Explanations are given in as intuitive a way as possible, sometimes with unavoidable lengthy mathematical expressions which can be skipped by readers with insufficient time. In fact, we have put more emphasis on the conceptual aspects than on mathematical derivations. The author hopes that this method of presentation will not only lead readers to a deeper understanding but will also encourage them to further creative thinking.

Yutaka Toyozawa

Acknowledgments

The author is much indebted to the Research Institute for Fundamental Physics of Kyoto University, the Institute for Solid State Physics of the University of Tokyo and the Physics Department of Chuo University, where the basic ideas underlying this book were developed. His particular thanks are due to the late Professor R. Kubo for many useful suggestions and continual encouragement. It is his great pleasure to acknowledge helpful conversations and interactions with many colleagues in the field: Professors Masaharu Inoue, K. S. Song, E. Hanamura, K. Cho, Y. Onodera, H. Sumi, A. Sumi, A. Kotani, K. Nasu, V. Hizhnyakov, Y. Kayanuma, Y. Shinozuka, S. Abe, M. Schreiber and many others not mentioned here. The author has been happy to learn a great deal from experimental physicists, including Professors H. Kanzaki, K. Kobayashi, T. Koda, T. Goto and K. Kan'no, who revealed many of the intriguing mechanisms of nature.

The author acknowledges with thanks the excellent cooperation he received from the staff of Cambridge University Press, in particular to Dr. Simon Capelin, the publishing director for physical sciences, for his encouragement and generous consideration of the time needed for the completion of the manuscript, and to Dr. Brian Watts, the copy editor, for laboriously converting the incomplete manuscript into a form useful for a broader area of readers. Finally, his thanks are due to Asako Toyozawa for her encouragement and help throughout this endeavor.

Yutaka Toyozawa

Principal symbols

A	Einstein's A coefficient; amplitude (Chapter 5)
A_a	electron affinity of an acceptor
$A(\mathbf{r})$	vector potential
a	microscopic length, such as orbital radius
a_{ex}	exciton radius
$a(\mathbf{r})$	Wannier function
a_0	lattice constant
a, a^\dagger	annihilation, creation operator of a fermion
B	Einstein's B coefficient; half-width of a band
\mathbf{B}	magnetic flux density
b	reciprocal lattice
b, b^\dagger	annihilation, creation operator of a boson
C	elastic constant
c	light velocity
c_ℓ	longitudinal sound velocity
c_0	light velocity in vacuum
D	amplitude of local energy fluctuation
\mathbf{D}	electric flux density
d	dimensionality
E	energy
E_0	energy of zero-phonon line
E_a	energy of peak (center) of absorption band
E_e	energy of peak (center) of emission band
E_R	relaxation energy
\mathbf{E}	electric field
$\mathbf{e}(\mathbf{r})$	microscopic electric field (Section 6.3)
e	absolute value of the electronic charge

e, e^*	effective charge of ion (Chapter 5)
\mathbf{e}	unit vector for polarization of light
$F(E), F_a(E)$	absorption spectrum
$F_e(E)$	emission spectrum
$f(t)$	generating function
f	oscillator strength; force constant
\mathbf{f}	force density
g	degree of degeneracy; coupling constant
g_s, g_ℓ	short- and long-range coupling constants,
H, h	Hamiltonian
\mathbf{H}	magnetic field
I_d	ionization energy of a donor
i, j, k	(suffix) rectangular component
i, j, k	(suffix) particle numbering
$\hat{i}, \hat{j}, \hat{k}$	unit vectors along three rectangular axes
\mathbf{J}	electric current density
$\mathbf{j}(\mathbf{r})$	operator for microscopic electric current density
K	kinetic energy
$\mathbf{K}, \mathbf{k}, \mathbf{q}$	wave vector
k_B	Boltzmann constant
L	macroscopic length
L_T	thermal energy density (Chapter 1)
ℓ	(suffix) longitudinal; long range
l	azimuthal quantum number (ℓ in Chapter 12 only)
\mathbf{l}	orbital angular momentum
ℓ, m, n	state- or site-numbering; directional cosines
M	atomic mass
M_r	atomic reduced mass (Chapter 5)
m	electron mass; magnetic quantum number
m_e	electron effective mass
m_h	hole effective mass
$m_{\text{ex}} (= m_e + m_h)$	exciton mass
$m_r [(m_e^{-1} + m_h^{-1})^{-1}]$	exciton reduced mass
N	total number
N_0	number density
n	principal quantum number
$n (\equiv b^\dagger b, \equiv a^\dagger a)$	number operator (for boson, or fermion)
$n(\omega)$	refractive index
$O(\dots)$	of the order of . . .

$p, \mathbf{p}, \mathbf{P}$	momentum of a particle
$\mathbf{P}(\mathbf{r})$	electric polarization density
$p_e [\equiv m_h/(m_e + m_h)]$; $p_h [\equiv m_e/(m_e + m_h)]$	fractional distances of an electron and a hole from their center-of-mass coordinate
Q	point charge (Chapter 6)
q, \mathbf{q}, Q	position coordinate of a particle
R_y	Rydberg energy
\mathbf{r}	position in space
s	(suffix) short range
s	mode index; spin; steepness index (Section 10.6)
s	spin angular momentum
$s(\omega)$	spectral coupling (coupling function)
S	coupling strength; total spin
T	absolute temperature
t	(suffix) transverse; triplet
t	time; interatomic transfer energy
U, V	potential energy
V	macroscopic volume
v	velocity; potential energy
W	rate of a process (Section 6.6)
$W_{g,e}(q)$	adiabatic potential of ground or excited state
$w(\omega)$	spectral density of energy (Chapter 3)
$w(q); w_\ell$	Boltzmann distribution
x, y, z	rectangular coordinates
Z	partition function
α	polaron coupling constant
$\beta (\equiv 1/k_B T)$	reciprocal temperature in units of reciprocal energy
γ	decay rate ($= 2 \Gamma/\hbar$); non-adiabaticity parameter (Section 9.4)
2Γ	level width (see Σ)
Δ	level shift (see Σ); displacement (Chapter 4); exchange energy (Section 8.5); ℓ -t splitting (Section 8.6); exciton energy difference in mixed crystal (Section 8.7)
$\Delta(\mathbf{r})$	dilation field
$\nabla \equiv (\partial/\partial x, \partial/\partial y,$ $\partial/\partial z)$	gradient operator
ε	one-electron energy,

ε_g	bandgap energy
ε_μ	normally occupied level (Chapters 8, 12)
ε_v	normally unoccupied level (Chapters 8, 12)
$\epsilon; \epsilon(\omega)$	dielectric constant; dielectric function
$\epsilon_0 \equiv \epsilon(\infty)$	dielectric constant of vacuum
ϵ_e	electronic dielectric constant
$\epsilon_s \equiv \epsilon(0)$	static dielectric constant
λ	wavelength; mode of photons specified by (\mathbf{k} , \mathbf{e}) (Chapter 3); spin–orbit coupling constant; quantum number for internal states of an exciton
$\Lambda(\mathbf{r})$	locator (Chapter 6)
μ	magnetic permeability
μ_0	magnetic permeability of vacuum
$\boldsymbol{\mu}$	electric dipole moment
ν	frequency; number of nearest neighbors; number of p-holes in an atom (Section 13.1)
ρ	charge density; density matrix (Chapter 2); mass density (Chapter 5); number density (Chapter 6)
$\rho(E)$	density of states (per energy)
\prod	product
\sum	summation
$\Sigma (= \Delta + i\Gamma)$	self-energy
σ	electric conductivity; steepness coefficient (Urbach rule); number of atoms in unit cell
τ	relaxation time; lifetime; other characteristic time
χ	electric susceptibility
$\Omega, \omega (\equiv 2\pi\nu)$	angular frequency; energy (in Sections 11.2–11.7 only)
ω_t, ω_ℓ	angular frequencies of transverse, longitudinal optical modes lattice vibrations
$(\hbar)\omega_{nm}$	(with two suffixes) $\equiv E_n - E_m$ (without \hbar in Sections 11.2–11.7)
$d\Omega$	differential solid angle
Ξ	deformation potential constant
$\boldsymbol{\xi}(\boldsymbol{\eta}, \boldsymbol{\zeta})$	displacement of an atom (Chapter 5)
ζ	chemical shift (Section 12.2)
Φ, Ψ, ϕ, ψ	wave function
$\phi(\mathbf{r}), \Phi(\mathbf{r})$	scalar potential
$\text{Tr}(A)$	trace: sum of diagonal elements
$[A, B]_- \equiv AB - BA$	commutator

$[A, B]_+ \equiv AB + BA$	anticommutator
$\langle A \rangle$	quantum-mechanical expectation value
$\langle\langle A \rangle\rangle$	thermodynamic average
$\overline{\overline{A}}$	tensor of the second rank; renormalized quantity (in Chapter 10 only)
$\mathbf{A} \cdot \mathbf{B}$	scalar product
$\mathbf{A} \times \mathbf{B}$	vector product

1

Radiation field

1.1 Maxwell equations for electromagnetic field

Early in the nineteenth century, people already realized that light was better described by a wave model than by a corpuscular model on the basis of experimental confirmation of the interference effect. However, they were not yet aware of an entity that could oscillate and propagate in a vacuum, unlike the case of a sound wave which they knew to be a compressional wave of air. The answer came from an apparently different area: later in that century, Maxwell formulated the experimental laws relating the spatial and temporal changes of electric and magnetic fields into a set of simultaneous equations which are now called by his name. He found the set to have a solution describing an electromagnetic wave which propagates in a vacuum, and this wave was identified with light since the predicted velocity of the former exactly agreed with that of the latter already known at that time.

Maxwell's equations for the electric field \mathbf{E} and magnetic field \mathbf{H} are given as a set of the following four equations^{1,2}

$$\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t, \quad (1.1.1)$$

$$\nabla \times \mathbf{H} = \partial \mathbf{D} / \partial t + \mathbf{J}, \quad (1.1.2)$$

$$\nabla \cdot \mathbf{D} = \rho, \quad (1.1.3)$$

$$\nabla \cdot \mathbf{B} = 0. \quad (1.1.4)$$

Here $\nabla \equiv (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ denotes a vector differential operator, \times the vector product and \cdot the scalar product. Therefore, $\nabla \cdot \mathbf{B}$ represents $\text{div} \mathbf{B} \equiv (\partial/\partial x)B_x + \dots$ and $\nabla \times \mathbf{E}$ represents $\text{rot} \mathbf{E} \equiv [(\partial/\partial y)E_z - (\partial/\partial z)E_y, \dots, \dots]$.

For the *electric flux density* \mathbf{D} and *magnetic flux density* \mathbf{B} , we have also linear relations which are exact in a vacuum and hold approximately in many ordinary

(neither ferroelectric nor ferromagnetic) materials for weak fields:

$$\mathbf{D} = \epsilon \mathbf{E}, \quad (1.1.5)$$

$$\mathbf{B} = \mu \mathbf{H}. \quad (1.1.6)$$

The *dielectric constant* ϵ and the *magnetic permeability* μ are material constants, while in a vacuum they are universal constants written as ϵ_0 and μ_0 . The most recent value of ϵ_0 is $8.854\,187\,817 \times 10^{-12} \text{ F m}^{-1}$, while $\mu_0 = 4\pi \times 10^{-7} \text{ N A}^{-2}$ by definition.

Matter with a *charge density* ρ is subject to the *Lorentz force* with density \mathbf{f} and velocity \mathbf{v} which is given by

$$\mathbf{f} = \rho(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (1.1.7)$$

This equation should be added to (1.1.1–1.1.4) for a consistent description of interacting particles and fields satisfying the conservation laws of energy and momentum, as will be seen below.

Taking the divergence of (1.1.2) and using (1.1.3), one gets the *continuity equation* describing the conservation of charge through the *electric current* \mathbf{J} :

$$\partial \rho / \partial t + \nabla \cdot \mathbf{J} = 0. \quad (1.1.8)$$

Under the proportionality (1.1.5, 1.1.6), the energy density of the fields is given by

$$U \equiv (\mathbf{D} \cdot \mathbf{E} + \mathbf{B} \cdot \mathbf{H})/2, \quad (1.1.9)$$

as is seen by integrating it over the volume V and taking its time derivative:

$$d/dt \int_V U dV = \int_V (\partial \mathbf{D} / \partial t \cdot \mathbf{E} + \partial \mathbf{B} / \partial t \cdot \mathbf{H}) dV.$$

With the use of eqs. (1.1.1, 1.1.2), the identity: $\nabla \cdot (\mathbf{E} \times \mathbf{H}) = \mathbf{H} \cdot \nabla \times \mathbf{E} - \mathbf{E} \cdot \nabla \times \mathbf{H}$ and the partial integration to be balanced by the surface integral, one finally obtains the energy conservation relation:

$$d/dt \int_V U dV = - \int_V (\mathbf{J} \cdot \mathbf{E}) dV - \int_s S_n ds. \quad (1.1.10)$$

Here the first integral on the right hand side (r.h.s.) represents the energy dissipated as *Joule's heat* and the second the energy lost through the surface s (suffix n denotes the component outward-normal to the surface element ds) by the flow:

$$\mathbf{S} \equiv \mathbf{E} \times \mathbf{H}, \quad (1.1.11)$$

which is called the *Poynting vector*.

1.2 Electromagnetic wave

Under the absence of charge ρ and current \mathbf{J} , (1.1.1–1.1.6) reduce to homogeneous equations for the fields. Putting (1.1.5, 1.1.6) into (1.1.1, 1.1.2) and making use of the identity: $\nabla \times \nabla \times \mathbf{E} = \nabla(\nabla \cdot \mathbf{E}) - (\nabla^2)\mathbf{E}$, where $\nabla^2 \equiv (\nabla \cdot \nabla) = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$, one obtains

$$\nabla^2 \mathbf{E} - \epsilon\mu \partial^2 \mathbf{E} / \partial t^2 = 0 \quad (1.2.1)$$

and a similar equation for \mathbf{B} . Equation (1.2.1) has a solution describing a plane wave

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)] \quad (1.2.2)$$

which propagates with *wave vector* \mathbf{k} ($k = 2\pi/\lambda$, where λ is the *wavelength*), *angular frequency* $\omega (= 2\pi\nu$, where ν is the frequency) and velocity $c = \omega/k$ given by

$$c = (\epsilon\mu)^{-1/2}. \quad (1.2.3)$$

In a vacuum, (1.2.3) reduces to $c_0 = (\epsilon_0\mu_0)^{-1/2} (= 2.997\,924\,58 \times 10^8 \text{ m s}^{-1}$, the most recent value), which was found to agree with the velocity of light as mentioned above. The \mathbf{E} wave (1.2.2) and the corresponding \mathbf{B} wave of the same form satisfy

$$\mathbf{k} \cdot \mathbf{E}_0 = 0, \quad \mathbf{k} \cdot \mathbf{B}_0 = 0, \quad \mathbf{B}_0 = i(\mathbf{k}/\omega) \times \mathbf{E}_0 \quad (1.2.4)$$

as is seen from (1.1.1, 1.1.3 and 1.1.4). Namely, we have a *transverse* wave with \mathbf{k} , \mathbf{E}_0 and \mathbf{B}_0 forming a right-handed system with \mathbf{B}_0 lagging behind \mathbf{E}_0 by phase difference $\pi/2$. From (1.1.9, 1.1.11) and (1.2.3, 1.2.4) one finds that $\mathbf{S} = c(\mathbf{k}/k)U$, namely that the energy density U is conveyed towards the direction of \mathbf{k} with velocity c .

Similarly to the energy conservation, one can derive the momentum conservation relation from (1.1.7) and the Maxwell equations. One can see from this relation that the electromagnetic wave has momentum density U/c in the direction of propagation \mathbf{k}/k , although we do not give the derivation here.

As the electromagnetic wave passes from a vacuum into a material, the velocity c_0 and hence the wavelength $\lambda_0 = c_0/\nu$ decrease to c and $\lambda = c/\nu$, respectively, since ν does not change due to the continuity of the field at the surface. The ratio

$$n \equiv c_0/c = (\epsilon\mu/\epsilon_0\mu_0)^{1/2}, \quad (1.2.5)$$

which is called *refractive index* of the material, relates the refraction angle θ to the incidence angle θ_0 through the law of refraction: $\sin \theta = (1/n) \times \sin \theta_0$.

1.3 Canonical equations of motion for electromagnetic waves

The wave equation (1.2.1) is a differential equation of second order in time t which is similar to the Newtonian equation of motion of a particle, whereas the original simultaneous equations (1.1.1, 1.1.2) for \mathbf{E} and \mathbf{B} are of first order in t , similar to the *canonical equations* of motion for the position coordinate q and the momentum p . The canonical form is more convenient, especially when one constructs quantum mechanical and quantum electrodynamical equations from the classical ones.

It is convenient to introduce the *vector* and *scalar potentials*, $\mathbf{A}(\mathbf{r}, t)$ and $\phi(\mathbf{r}, t)$, in place of the electromagnetic fields, $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{B}(\mathbf{r}, t)$. According to one theorem in vector analysis, the divergence-free field \mathbf{B} (eq. (1.1.4)) can be written as a rotation of the vector potential \mathbf{A} :

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (1.3.1)$$

Putting this into (1.1.1), one obtains

$$\nabla \times (\mathbf{E} + \partial \mathbf{A} / \partial t) = 0. \quad (1.3.2)$$

According to another theorem, the rotation-free field as given in $(\cdot \cdot \cdot)$ of (1.3.2) can be written as a gradient of the scalar potential ϕ :

$$\mathbf{E} + \partial \mathbf{A} / \partial t = -\nabla \phi. \quad (1.3.3)$$

A *general* solution of the set of inhomogeneous equations (1.1.1–1.1.4) for ϕ and \mathbf{A} is given as the sum of a *particular* solution of the set and a general solution of the homogeneous equations (obtained by putting $\mathbf{J} = 0$ and $\rho = 0$ in (1.1.1–1.1.4)). One particular solution which is well known is the *retarded potentials* $\phi(\mathbf{r}, t)$ and $\mathbf{A}(\mathbf{r}, t)$ due to the charge and current densities ρ and \mathbf{J} as their respective sources at the spacetime point $(\mathbf{r}', t' = t - |\mathbf{r} - \mathbf{r}'|/c)$, with the common integration kernel (over \mathbf{r}') being coulombic: $|\mathbf{r} - \mathbf{r}'|^{-1}$, as it should be. This solution plays an especially important role in studying how the oscillating charge and current as sources give rise to electromagnetic waves at a remote location. We will not give the derivation here, leaving it to the standard text books.^{1,2}

We are here concerned with a general solution of the homogeneous equations, namely a superposition of the electromagnetic waves which are given by

$$\mathbf{A}(\mathbf{r}, t) = \mathbf{A}_0 \exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)], \quad \mathbf{A}_0 = \mathbf{E}_0 / i\omega = (ik^{-2})\mathbf{k} \times \mathbf{B}_0, \quad (1.3.4)$$

as can be confirmed by eqs. (1.3.1, 1.3.3). They also satisfy the wave equation of the form (1.2.1). The wave vector \mathbf{k} can take arbitrary values in infinitely extended space. A more realistic situation is a confined space such as a cavity. For the

sake of simplicity, let us consider a cube of side L and impose a cyclic boundary condition

$$\mathbf{A}(\mathbf{r} + L\hat{i}, t) = \mathbf{A}(\mathbf{r} + L\hat{j}, t) = \mathbf{A}(\mathbf{r} + L\hat{k}, t) = \mathbf{A}(\mathbf{r}, t) \quad (1.3.5)$$

where \hat{i}, \hat{j} and \hat{k} are unit vectors along the three sides of the cube. Then \mathbf{k} in (1.3.4) takes only discrete values:

$$k_x = 2\pi n_x/L, k_y = 2\pi n_y/L, k_z = 2\pi n_z/L \quad (1.3.6)$$

where n_x, n_y and n_z can take independently all integer values, inclusive of negative values and zero.[†] Then the number of possible modes with k_x within a small interval dk_x is given by $dn_x = (L/2\pi)dk_x$ and similarly for y and z components. Since the electromagnetic wave is a transverse wave and there are two possible directions of polarization (defined to be along \mathbf{E}_0) for a given wave vector \mathbf{k} , the number per volume of possible modes contained within the spherical shell between k and $k + dk$ is given by

$$2dn/V = 2 \times 4\pi k^2 dk / 8\pi^3 = 2 \times 4\pi \rho(\omega) d\omega, \\ \rho(\omega) \equiv \omega^2 / 8\pi^3 c_0^3. \quad (1.3.7)$$

The set of all $u_{\mathbf{k}}(\mathbf{r}) = L^{-3/2} \exp(i\mathbf{k} \cdot \mathbf{r})$ with \mathbf{k} given by (1.3.6) forms an orthonormal complete set such that $(u_{\mathbf{k}}, u_{\mathbf{k}'}) \equiv \int d\mathbf{r} u_{\mathbf{k}}^*(\mathbf{r}) u_{\mathbf{k}'}(\mathbf{r}) = \delta_{\mathbf{k}, \mathbf{k}'}$. However, in order to expand the vector potential $\mathbf{A}(\mathbf{r}, t)$, one has to prepare the set of vector basis functions

$$\mathbf{A}_{\mathbf{k}j}(\mathbf{r}) = \mathbf{e}_{\mathbf{k}j} u_{\mathbf{k}}(\mathbf{r}) \quad (j = 1, 2) \quad (1.3.8)$$

with the use of the two possible directions of polarization $\mathbf{e}_{\mathbf{k}j}$ for each \mathbf{k} which satisfy $\mathbf{e}_{\mathbf{k}j} \perp \mathbf{k}$ (transverse wave!) and $\mathbf{e}_{\mathbf{k}j} \cdot \mathbf{e}_{\mathbf{k}j'} = \delta_{jj'}$. Denoting the set $(\mathbf{k}j)$ simply by κ , one can see that (1.3.8) form an orthonormal set in the following sense:

$$\int d\mathbf{r} \mathbf{A}_{\kappa}^*(\mathbf{r}) \cdot \mathbf{A}_{\kappa'}(\mathbf{r}) = \delta_{\kappa\kappa'}. \quad (1.3.9)$$

Then one can expand

$$\mathbf{A}(\mathbf{r}, t) = (4\epsilon)^{-1/2} \sum_{\kappa} [q_{\kappa}(t) \mathbf{A}_{\kappa}(\mathbf{r}) + q_{\kappa}^*(t) \mathbf{A}_{\kappa}^*(\mathbf{r})]. \quad (1.3.10)$$

Although the first term in $[\dots]$ is sufficient as an expansion, the addition of the second term automatically assures that the three components of the vector $\mathbf{A}(\mathbf{r}, t)$

[†] A more realistic electromagnetic boundary condition for the real-valued waves of the form $\sin(\mathbf{k} \cdot \mathbf{r}) \sin(\omega t)$ may be that the waves vanish at the boundary surface, which gives possible k values without the factor 2 on the right hand sides of (1.3.6) but confines them to positive values. The two effects cancel out, keeping the number of possible modes within the interval $(k, k + dk)$, as given by (1.3.7), unchanged. Other possible shapes of the cavity do not change (1.3.7) as long as we are concerned with the majority of k values which are much greater than $1/L$.

are *real* quantities for any t . (The prefactor with ϵ , the dielectric constant, is for the normalization of energy, as will be seen later.) It is convenient to introduce *real*-valued dynamical coordinates

$$Q_\kappa(t) \equiv [q_\kappa(t) + q_\kappa^*(t)]/2 \quad (1.3.11)$$

instead of the complex-valued q_κ s. In order that each term in (1.3.10) represents a running wave as in eq. (1.2.2), $q_\kappa(t)$ should vary as $\exp(-i\omega_\kappa t)$ where $\omega_\kappa \equiv \omega_k$. The time derivative of (1.3.11) is then given by

$$dQ_\kappa/dt = -i\omega_\kappa(q_\kappa - q_\kappa^*)/2. \quad (1.3.12)$$

By introducing the momentum $P_\kappa \equiv dQ_\kappa/dt$ as an independent dynamical variable, one can rewrite the complex q_κ , with the use of (1.3.11, 1.3.12), as

$$q_\kappa = Q_\kappa + (i/\omega_\kappa)P_\kappa. \quad (1.3.13)$$

Now the energy density (1.1.9) can be rewritten as

$$U = (\epsilon/2)[(\partial \mathbf{A}/\partial t)^2 + c^2(\nabla \times \mathbf{A})^2]$$

with the use of eqs. (1.1.5, 1.1.6) and (1.3.1, 1.3.3). Putting (1.3.10, 1.3.13) into this expression and integrating over the volume which is the cube, one obtains

$$H \equiv \int U d\mathbf{r} = \sum_\kappa (P_\kappa^2 + \omega_\kappa^2 Q_\kappa^2)/2 \quad (1.3.14)$$

with the use of (1.3.9). The total energy of the transverse electromagnetic field is thus reduced to a Hamiltonian H of an assembly of harmonic oscillators indexed with κ . In fact, the canonical equations of motion: $dQ_\kappa/dt = \partial H/\partial P_\kappa$ and $dP_\kappa/dt = -\partial H/\partial Q_\kappa$ give an harmonic oscillation with angular frequency ω_κ .

The above procedure may seem to be artificial because in defining P_κ by (1.3.12) we have made use of a solution yet to be obtained. However, this artifice is common to the canonical formalism in which the formula defining the momentum $P \equiv mdQ/dt$ with mass m is to be derived from one of the canonical equations of motion: $dQ/dt = \partial H/\partial P$ ($= P/m$ for a general motion which is not harmonic). Introduction of P as an independent variable has the merit of lowering the differential equations of motion to first order in time as compared to the Newtonian equations of motion which are of second order. This facilitates the time-integration a great deal, so as to more than cover the demerit of doubling the number of independent variables. The canonical formalism turns out to be more natural in statistical mechanics in which P and Q are independent variables, and to be indispensable for quantum mechanics in which P and Q are mutually conjugate variables subject to the uncertainty relation.

1.4 Thermal radiation field

We are now in a position to describe that experimental fact whose deviation from the predictions of classical physics provided a clue to the discovery of quantum mechanics. It is well known, even in daily experience, that the glowing radiation emitted by a hot body shifts from red to yellow, namely toward shorter wavelengths, as the temperature increases. The observed spectral distribution of this radiation in an idealized situation, namely in a cavity which is in thermal equilibrium with its wall, depends only on the temperature but not on the material of the wall, as shown schematically by the solid lines in Fig. 1.1. The broken lines are the spectra predicted by the classical theory available at that time, late in the 19th century when the experiments were done. One finds that the theory deviates from the observation significantly in the short-wavelength region.

Let us trace this theoretical prediction which was based on the classical mechanics and statistical mechanics applied to an assembly of harmonic oscillators representing the radiation field as mentioned in the preceding section. A harmonic oscillator with Hamiltonian given by $H(Q, P) = P^2/2 + \omega^2 Q^2/2$, which is in thermal equilibrium at temperature T , obeys the Boltzmann distribution: the probability of finding the system within a small region of the phase space (Q, P) is proportional to

$$\exp[-\beta H(Q, P)]dQdP, \text{ with } \beta \equiv 1/k_B T \quad (1.4.1)$$

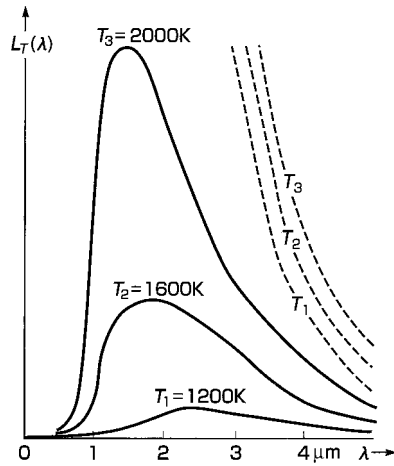


Fig. 1.1 The spectral distribution of energy density, L_T , of thermal radiation field at various temperatures. Solid lines are experimental results (shown schematically) which can be fitted by the quantum hypothesis of Planck. Broken lines are the predictions of classical field theory.

where k_B is the Boltzmann constant. Consider the region of the phase space which is surrounded by an ellipse given by $H(Q, P) = E$. Its area is given by

$$s(E) = \pi(2E/\omega^2)^{1/2}(2E)^{1/2} = E/\nu \quad (1.4.2)$$

where $\nu = \omega/2\pi$ is the frequency. Integrating (1.4.1) over the elliptic shell between E and $E + dE$ whose area is given by dE/ν , one obtains the Boltzmann distribution for the energy: $\exp(-\beta E)dE/\nu$. The statistical average of the energy is then given by

$$\begin{aligned} \langle\langle E \rangle\rangle &= \int_0^\infty E \exp(-\beta E)d(E/\nu) / \int_0^\infty \exp(-\beta E)d(E/\nu) \\ &= \beta^{-1} = k_B T \end{aligned} \quad (1.4.3)$$

which does not depend on the frequency ν of the oscillator for the electromagnetic wave concerned.

Let us rewrite (1.3.7), the number of normal modes of electromagnetic wave within a small interval dk of wave number, into that within an interval $d\lambda$ of wavelength, making use of the relation $k = 2\pi/\lambda$. The number per unit volume of the cavity space is then given by $8\pi\lambda^{-4}d\lambda$. Multiplying this by the energy per mode (1.4.3), one obtains the energy density per λ of the thermal radiation field

$$L_T(\lambda) = 8\pi k_B T \lambda^{-4} \quad (1.4.4)$$

which diverges at shorter wavelengths, as shown by the broken lines in Fig. 1.1.

In 1900, Planck put forth a working hypothesis,³ by which he could *provisionally* evade the discrepancy of the theoretical prediction from the experimental result; that concerned with the mechanics, not with the statistical distribution. In classical mechanics, one tacitly assumes that the energy of a particle in motion can take continuous values as exemplified by the continuous distribution shown in the integral of (1.4.3). For a harmonic oscillator with frequency ν , he *tentatively replaced* the continuous values by the following set of *discrete* values:

$$E_n = nh\nu \quad (n = 0, 1, 2, \dots). \quad (1.4.5)$$

Here, h is an empirical constant to be determined later. The integral in (1.4.3) is then to be replaced by a summation, with the following result:

$$\begin{aligned} \langle\langle E \rangle\rangle &= \sum_{n=0}^{\infty} nh\nu \exp(-\beta nh\nu) / \sum_{n=0}^{\infty} \exp(-\beta nh\nu) \\ &= h\nu / [\exp(\beta h\nu) - 1] = h\nu \langle\langle n \rangle\rangle. \end{aligned} \quad (1.4.6)$$

(Note: the summation in the denominator gives $f(\beta) = [1 - \exp(-\beta h\nu)]^{-1}$ so that the numerator can be written as $-df(\beta)/d\beta$.) Hence, (1.4.4) was replaced by

$$L_T(\lambda) = 8\pi\lambda^{-5}hc_0 / [\exp(hc_0/k_B T\lambda) - 1]. \quad (1.4.7)$$

Planck could reproduce the experimental spectra shown schematically in Fig. 1.1. at all temperature by the new formula (1.4.7) by choosing an appropriate value for h , the only adjustable parameter. In spite of this brilliant success, it took some time before people realized the physical meaning of the working hypothesis (1.4.5) which seemed so absurd from the viewpoint of classical physics. However, the constant h introduced by him as an empirical parameter played a pivotal role in the groping effort for new physical principles until quantum mechanics was discovered a quarter of a century later. And in fact Planck's hypothesis, (1.4.5), was successfully derived subsequently when quantum mechanics was established, as will be described in Sections 2.2 and 2.3. Planck's constant h is now a universal constant governing the entire quantum world. Its value is $6.626\,068\,76 \times 10^{-34}$ J s according to a recent measurement.

Returning to the problem of thermal radiation, we find that the new formula (1.4.7) reduces to the classical one (1.4.4) in the limit of high temperature or long wavelength, namely when $k_B T \lambda / hc_0 = k_B T / h\nu \gg 1$. The spectral maximum of the quantal expression (1.4.7) appears at λ_{\max} which is related to T by

$$\lambda_{\max} T = (hc_0/4.966k_B) = 0.290 \text{ cm K}. \quad (1.4.8)$$

This reproduces Wien's displacement law, the empirical law which had been found before Planck put forth his hypothesis.

In contrast to traditional spectroscopy in which λ has been chosen as the variable, it is more convenient to choose an angular frequency $\omega = 2\pi c_0/\lambda$. Then the energy density in ω -space, defined by $W_T(\omega)d\omega = L_T(\lambda)d\lambda$, is given by

$$W_T(\omega) = (\hbar\omega^3/\pi^2c_0^3)\langle\langle n \rangle\rangle = (\hbar\omega^3/\pi^2c_0^3)/[\exp(\beta\hbar\omega) - 1], \quad (1.4.9)$$

where we have defined $\hbar \equiv h/2\pi$ which is nowadays more frequently used than h . It takes its maximum value at

$$\hbar\omega_{\max} = 2.821k_B T \quad (1.4.10)$$

which is different from the photon energy corresponding to the λ_{\max} of (1.4.8), namely $hc_0/\lambda_{\max} = 4.966k_B T$.

2

Quantum mechanics

2.1 Elements of quantum mechanics

Quantum mechanics was discovered in 1925 through groping efforts to compromise two apparently contradictory pictures on the fundamental entities in nature. One was the *wave* picture for light which was later extended to matter by de Broglie, another was the *corpuscular* picture of matter which was later extended to light by Einstein. Schrödinger's wave equation came as a natural development of the first stream, while Heisenberg's matrix mechanics was presented as a unique proposal from the second stream. In spite of completely different appearances, the two theories proved, within a couple of years after their discoveries, to be equivalent. This is a most beautiful example that the physical reality exists independent of the mathematical framework formulated for its description.

In this chapter, we will give a very brief review of the principles of quantum mechanics,¹⁻³ mainly with the harmonic oscillator as a model system for the following reasons. The first is historical: the electromagnetic wave, whose interaction with matter is the subject of this book, is a harmonic oscillator, a system which was for the first time subject to “quantization”, thus opening a way to the discovery of quantum mechanics. The second is technical: the harmonic oscillator is one of very few examples of analytically soluble problems in quantum mechanics. The third is pedagogical: the harmonic oscillator is a system best suited for realization of the equivalence of the two different pictures mentioned above and hence for a deeper understanding of the principles of quantum mechanics. Finally, the fourth is practical and applies particularly to matter: in this book we will deal with a variety of elementary excitations in solids, such as phonons, excitons and plasmons which are approximately harmonic oscillators.

According to quantum mechanics, the *state* of a physical system is described by a *wave function* ψ_t which is a *complex* quantity and varies with time t following

the Schrödinger equation:

$$i\hbar(d/dt)\psi_t = H\psi_t. \quad (2.1.1)$$

In quantum mechanics, the dynamical variables A , such as position q , momentum p and the Hamiltonian H which is their functional – typically of the form $H(q, p) = p^2/2m + V(q)$ – are *linear operators* (instead of ordinary numbers as in classical mechanics) satisfying, by definition,

$$A(c_1\phi_1 + c_2\phi_2) = c_1A\phi_1 + c_2A\phi_2 \quad (2.1.2)$$

for any wave function ϕ and complex number c . Operators, which transform a wave function to another wave function as already implied in the above equations, can be treated like ordinary numbers in their addition and subtraction, but *not* in their *multiplication*. Namely, for arbitrary operators A and B , $A(B\phi) \neq B(A\phi)$ in general. Such quantities which do not in general follow the commutation relation $AB = BA$ are called *q-numbers* (q is an abbreviation of “quantal”) while the ordinary numbers which follow the commutation rule are called *c-numbers* (c for “classical”). If the equality holds for any ϕ , namely if $AB = BA$ (product of operators is defined as $AB\phi \equiv A(B\phi)$), one states that A *commutes* with B or vice versa.

The basic postulate of quantum mechanics is that position q and momentum p satisfy the *commutation relation*:

$$[q, p]_- \equiv qp - pq = i\hbar. \quad (2.1.3)$$

The quantity $[]$ on the left hand side (l.h.s.) is called the *commutator*. In the “ q -diagonal representation” which is of most popular use, the position operator q is treated as an independent variable (a real-valued c-number in this representation only) to express $\phi(q)$ as a complex-valued function thereof, while p is a differential operator:

$$p = -i\hbar\partial/\partial q. \quad (2.1.4)$$

In fact, for an arbitrary function $\phi(q)$, we have

$$\{q(-i\hbar\partial/\partial q) - (-i\hbar\partial/\partial q)q\}\phi(q) = i\hbar\phi(q)$$

which is equivalent to the operator equation (2.1.3).

The *expectation value* of a physical quantity represented by an operator A , in the state ψ_t , is given by

$$\begin{aligned} \langle\langle A \rangle\rangle_t &= (\psi_t, A\psi_t)/(\psi_t, \psi_t) \\ &= \int dq \psi_t^*(q) A\psi_t(q) / \int dq \psi_t^*(q) \psi_t(q) \end{aligned} \quad (2.1.5)$$

with the use of the scalar product $(\phi, \psi) = (\psi, \phi)^*$ defined by

$$(\phi, \psi) \equiv \int dq \phi^*(q) \psi(q). \quad (2.1.6)$$

The wave function $c\psi(q)$, where c is an arbitrary complex number, represents the same physical state as $\psi(q)$ itself, as is implied by (2.1.5).

Putting the position operator q (the operator of simply multiplying q in our q -diagonal representation) for A of eq. (2.1.5) one finds that $\psi_t^*(q)\psi_t(q)dq = |\psi_t(q)|^2 dq$ represents the *probability* that the particle be found in $(q, q + dq)$ provided ψ_t is normalized: $(\psi_t, \psi_t) = 1$.

An operator A^\dagger which satisfies the relation

$$(A^\dagger \phi, \psi) = (\phi, A\psi) \quad (2.1.7)$$

for arbitrary wave functions ϕ, ψ is called the *conjugate operator* of A . The conjugate of a product of operators is equal to the inverted product of their conjugates: $(AB)^\dagger = B^\dagger A^\dagger$, as can be proved by repeated use of (2.1.7). An operator A which is conjugate to itself ($A^\dagger = A$) is called *self-conjugate* or an *Hermite operator*. The expectation value of an Hermite operator is real as is obvious from

$$(\psi, A\psi) \equiv (A^\dagger \psi, \psi) = (A\psi, \psi) = (\psi, A\psi)^*.$$

The momentum operator p defined by (2.1.4) is an Hermite operator as long as one confines oneself to wave functions which vanish at infinity or satisfy the cyclic boundary condition (1.3.5). In fact, by partial integration we have

$$\begin{aligned} (\phi, p\psi) &= \int_a^b dq \phi^*(q) (-i\hbar) \partial \psi(q) / \partial q \\ &= \int_a^b dq \{ i\hbar \partial \phi^*(q) / \partial q \} \psi(q) + (-i\hbar) \phi^*(q) \psi(q) \Big|_a^b \\ &= \int_a^b dq \{ -i\hbar \partial \phi(q) / \partial q \}^* \psi(q) = (p\phi, \psi) \end{aligned}$$

where (a, b) denotes $(-\infty, +\infty)$ or $(0, L)$. The Hermiticity of q is obvious, whence that of the functional $H(p, q) \equiv p^2/2m + V(q)$ follows immediately.

Multiplying (2.1.1) and its complex conjugate by $\psi_t^*(q)$ and $\psi_t(q)$, respectively, integrating them over q and subtracting one from the other, one obtains

$$i\hbar(d/dt)(\psi_t, \psi_t) = (\psi_t, H\psi_t) - (H\psi_t, \psi_t) = 0 \quad (2.1.8)$$

due to the Hermiticity of H . Namely, the *normalization* $(\psi_t, \psi_t) = 1$ holds at any time if it does at $t = 0$. One can then omit the denominator in (2.1.5).

Firstly, apply operator A to eq. (2.1.1) and multiply $\psi_t^*(q)$ from the left side. Secondly, take the complex conjugate of (2.1.1) and multiply $A\psi_t(q)$ from the

right side. Subtracting the former from the latter and integrating over q , one obtains

$$\begin{aligned} i\hbar(d/dt)\langle A \rangle_t &= i\hbar(d/dt)(\psi_t, A\psi_t) = (\psi_t, AH\psi_t) - (H\psi_t, A\psi_t) \\ &= (\psi_t, [A, H]_-\psi_t) \end{aligned} \quad (2.1.9)$$

where use has been made of the Hermiticity of H and of the commutator defined in (2.1.3). Putting H for A , one immediately obtains the energy conservation: $\langle H \rangle_t = \text{constant}$.

In the *Heisenberg picture*, the dynamical variable A varies with time according to the equation:

$$i\hbar(d/dt)A_t = [A_t, H]_- \quad (2.1.10)$$

while the state ψ does not depend on time. The expectation value of A_t , given by $\langle A_t \rangle \equiv (\psi, A_t\psi)$ as in the *Schrödinger picture* so far described, varies with time as

$$i\hbar(d/dt)\langle A_t \rangle = (\psi, [A_t, H]_-\psi). \quad (2.1.11)$$

Equations (2.1.9) and (2.1.11) describe the same reality in different ways, depending on whether one ascribes the evolution of the system to that of the wave function or to that of the dynamical variable. Equation (2.1.10) is called the *Heisenberg equation of motion*. We will hereafter mainly use the *Schrödinger picture*, resorting to the *Heisenberg picture* only when the latter is more convenient or appeals better to intuition.

If, for a given operator A , $A\psi$ represents the same state as ψ , namely, if

$$A\psi = a\psi, \quad (2.1.12)$$

the number a is called an *eigenvalue* of the operator A , and ψ is called the *eigenfunction* belonging to the eigenvalue a . For the latter reason, ψ in eq. (2.1.12) is often written as ψ_a . The eigenvalues of an Hermite operator are real as is proved in the same way as was done for the expectation value. For instance, the momentum operator (2.1.4) has plane waves $\psi_p(q) = \exp(ikq)$ as eigenfunctions each belonging to the eigenvalue $p = \hbar k$ where k can take any of the discrete values given by (1.3.6), which tend to continuous values in the limit of $L \rightarrow \infty$. Eigenfunctions ψ_a and $\psi_{a'}$ belonging to different eigenvalues a and a' are orthogonal to each other, $(\psi_a, \psi_{a'}) = 0$, as is readily proved: (i) consider the eigenvalue equations $A\psi_a = a\psi_a$ and $A\psi_{a'} = a'\psi_{a'}$, (ii) take the scalar product of $\psi_{a'}$ with the first equation and that of the second equation with ψ_a , (iii) subtract the equation resulting from the latter from that resulting from the former, (iv) note that the l.h.s. vanishes because of eq. (2.1.7) with $A^\dagger = A$ while the r.h.s. gives $(a - a')(\psi_{a'}, \psi_a)$, whence follows the orthogonality for $a \neq a'$. If there is more than one eigenfunction belonging to an eigenvalue, we can make them orthogonal to each other by taking an appropriate

set of their linear combinations: $\psi_{a_1}, \psi_{a_2}, \dots$. Renumerate all the suffixes by j , denote the corresponding eigenvalue by a_j , and normalize the eigenfunctions ψ_j . Then the eigenfunctions of an Hermite operator A form an orthonormal set: $(\psi_i, \psi_j) = \delta_{ij}$, as was the case with the $u_k(\mathbf{r})$ s of Section 1.3 which are the eigenfunctions of the momentum operator in three-dimensional space. Moreover the entire set of eigenfunctions (ψ_i) of an Hermite operator A forms a *complete set* in the sense that any wavefunction $\psi(q)$ can be expanded in terms of them as

$$\psi(q) = \sum_i c_i \psi_i(q), \text{ where } c_i = (\psi_i, \psi) \quad (2.1.13)$$

(as was also the case with the Fourier expansion of an arbitrary function in terms of the $u_k(\mathbf{r})$ s). The expectation value (2.1.5) can now be written as

$$\langle A \rangle = (\psi, A\psi) = \sum_j |c_j|^2 a_j \quad (2.1.14)$$

provided that ψ is normalized. It means that a physical quantity A can take only the values a_j (discrete and/or continuous depending on the operator A) and that the probability of finding the value a_j is given by $|c_j|^2$.

The eigenvalues E_j and eigenfunctions $\psi_j(q)$ of the Hamiltonian which satisfy

$$H(q, p)\psi_j(q) = E_j\psi_j(q) \quad (2.1.15)$$

play a particular role in considering the time dependence of the state ψ . Putting the linear combination of the form (2.1.13) into the Schrödinger equation (2.1.1) and multiplying ψ_j from the left and integrating, one obtains the equation $i\hbar dc_j/dt = E_j c_j$. Putting in the solution: $c_j(t) = c_j^0 \exp(-itE_j/\hbar)$, we obtain the general solution:

$$\psi(q, t) = \sum_j c_j^0 \exp(-itE_j/\hbar) \psi_j(q) \quad (2.1.16)$$

which will play an important role in studying the dynamical behavior of the system, as will be seen later.

In passing, it is to be noted that each term in (2.1.16) can be rewritten as

$$\exp(-itE_j/\hbar) \psi_j(q) = \exp(-itH/\hbar) \psi_j(q)$$

provided that one defines the function $f(H)$ of the operator H as an operator with eigenvalues $f(E_j)$ and eigenfunctions $\psi_j(q)$. Inserting this into (2.1.16), one can write

$$\psi(q, t) = \exp(-itH/\hbar) \psi(q, 0) \quad (2.1.17)$$

which is nothing but the result of formal integration of the Schrödinger equation (2.1.1) treating H as a c-number. The fact that such a procedure is justified under the

definition of the operator function $f(H)$ mentioned above will be used in Section 3.5 and later chapters.

Likewise, the Heisenberg equation of motion, (2.1.10), can be integrated as

$$A_t = \exp(iHt/\hbar)A_0 \exp(-iHt/\hbar) \quad (2.1.18)$$

(matrix elements: $(A_t)_{jk} = \exp(iE_j t/\hbar)A_0 \exp(-iE_k t/\hbar)$),

as is readily seen by differentiating (2.1.18) with respect to t :

$$\begin{aligned} dA_t/dt &= \exp(iHt/\hbar)(i/\hbar)[H, A_0]_ - \exp(-iHt/\hbar) \\ &= (i/\hbar)[H, \exp(iHt/\hbar)A_0 \exp(-iHt/\hbar)]_ - \end{aligned}$$

and confirming that the result agrees with (2.1.10).

2.2 Quantum mechanics of a harmonic oscillator

As was already mentioned in Section 1.4, it was with the harmonic oscillators of the electromagnetic waves in a thermal radiation field that the deviation from classical physics was first noticed and this finally opened a way to the discovery of quantum mechanics. Here we present the quantum mechanics of a harmonic oscillator, which is most instructive in realizing the nature of the difference between classical and quantum mechanics.

The Hamiltonian $H(q, p)$ of a harmonic oscillator is given by

$$H(q, p) = p^2/2m + m\omega^2 q^2/2 \quad (2.2.1)$$

where m is the mass of the particle and ω is the angular frequency. Let us first solve the *eigenvalue problem* of the energy:

$$\begin{aligned} H(q, -i\hbar d/dq)\psi(q) &\equiv \{-(\hbar^2/2m)(\partial/\partial q)^2 + (m\omega^2/2)q^2\}\psi(q) \\ &= E\psi(q), \end{aligned} \quad (2.2.2)$$

which is one of few analytically solvable eigenvalue problems in quantum mechanics. For simplicity, we introduce a dimensionless coordinate x with a unit of length a defined by

$$x \equiv q/a, \quad a \equiv (\hbar/m\omega)^{1/2}, \quad (2.2.3)$$

and a dimensionless operator b and its conjugate b^\dagger defined by:

$$b, b^\dagger \equiv 2^{-1/2}\{x \mp (d/dx)\} = 2^{-1/2}\{(1/a)q \pm i(a/\hbar)p\}, \quad (2.2.4)$$

in terms of which the original operators are given by

$$q = a \times 2^{-1/2}(b + b^\dagger), \quad p = (-i\hbar/a) \times 2^{-1/2}(b - b^\dagger). \quad (2.2.5)$$

The commutation relation (2.1.3) between the coordinate q and the momentum p , is now to be replaced by

$$[b, b^\dagger]_- \equiv bb^\dagger - b^\dagger b = 1. \quad (2.2.6)$$

With the use of an operator n defined by

$$n \equiv b^\dagger b = bb^\dagger - 1, \quad (2.2.7)$$

the energy operator given in (2.2.1) can be written in a compact form:

$$H = (bb^\dagger + b^\dagger b)\hbar\omega/2 = \left(n + \frac{1}{2}\right)\hbar\omega. \quad (2.2.8)$$

Comparing this with (2.2.2), we have only to solve the eigenvalue problem of n :

$$n\psi_\eta = \eta\psi_\eta. \quad (2.2.9)$$

Since the eigenvalue η is the expectation value in the corresponding eigenstate, one can rewrite

$$\eta = (\psi_\eta, n\psi_\eta) \equiv (\psi_\eta, b^\dagger b\psi_\eta) = (b\psi_\eta, b\psi_\eta)$$

under the normalization condition $(\psi_\eta, \psi_\eta) = 1$ (as will hereafter be assumed for any eigenfunctions unless otherwise mentioned). In the third equality we used the fact that b is the conjugate of b^\dagger . It is convenient to define ψ'_η by

$$\psi'_\eta \equiv b\psi_\eta, \quad (2.2.10)$$

with which one can further rewrite the above equation,

$$\eta = (\psi'_\eta, \psi'_\eta) = \int dq |\psi'_\eta(q)|^2 \geq 0 \quad (2.2.11)$$

Making use of eqs. (2.2.6, 2.2.7, and 2.2.10), one can write

$$n\psi'_\eta = (bb^\dagger - 1)b\psi_\eta = b(b^\dagger b - 1)\psi_\eta = (\eta - 1)\psi'_\eta. \quad (2.2.12)$$

Equation (2.2.12) indicates that ψ'_η is that eigenfunction (not normalized) of the operator n which belongs to eigenvalue $(\eta - 1)$. According to the inequality (2.2.11), there should be the smallest eigenvalue η_0 , which is either (i) > 0 or (ii) $= 0$. In case (i), $\eta_0 - 1$ is also an eigenvalue according to (2.2.12), since ψ'_η is not an identically vanishing function according to the inequality (2.2.11). This contradicts the definition of η_0 as the smallest eigenvalue, leaving only possibility (ii), namely $\eta_0 = 0$. In this way, we find that the eigenvalues of the operator n are given by

$$\eta = 0, 1, 2, \dots \quad (2.2.13)$$

Hereafter n will be called the *number operator* since its eigenvalues η are non-negative integers. One also finds from eq. (2.2.11) that $\psi_0' = 0$ when $\eta_0 = 0$, and so, from (2.2.10), that

$$b\psi_0 = 0. \quad (2.2.14)$$

While eq. (2.2.10) shows how to get the eigenfunction $\psi_{\eta-1}$ from ψ_η except for the normalization constant, one can obtain the eigenfunction $\psi_{\eta+1}$ from ψ_η by

$$c_{\eta+1}\psi_{\eta+1} = \psi_\eta'' \equiv b^\dagger\psi_\eta \quad (2.2.15)$$

except for the normalization constant c , as is confirmed as follows:

$$n\psi_\eta'' \equiv b^\dagger b b^\dagger\psi_\eta = b^\dagger(b^\dagger b + 1)\psi_\eta = b^\dagger(\eta + 1)\psi_\eta = (\eta + 1)\psi_\eta''.$$

The normalization factor c is obtained from

$$(\psi_\eta'', \psi_\eta'') = (b^\dagger\psi_\eta, b^\dagger\psi_\eta) = (bb^\dagger\psi_\eta, \psi_\eta) = (\eta + 1)(\psi_\eta, \psi_\eta)$$

as $c_{\eta+1} = (\eta + 1)^{1/2}$, and from (2.2.15) one obtains the normalized eigenfunction:

$$\psi_{\eta+1} = (\eta + 1)^{-1/2}b^\dagger\psi_\eta. \quad (2.2.16)$$

With the repeated use of this recursion formula, one obtains the set of orthonormal eigenfunctions:

$$\psi_\eta = (\eta!)^{-1/2}(b^\dagger)^\eta\psi_0 \quad (\eta = 0, 1, 2, \dots). \quad (2.2.17)$$

According to (2.1.13), an arbitrary wave function $\phi(q)$ can be expanded in terms of the orthonormal set $\psi_\eta(q)$ as

$$\phi(q) = \sum_\eta c_\eta \psi_\eta(q), \quad \text{where } c_\eta \equiv (\psi_\eta, \phi) \quad (2.2.18)$$

and the scalar product of two wave functions ϕ' and ϕ as

$$(\phi', \phi) = \sum_\eta c_\eta'^* c_\eta, \quad (2.2.19)$$

namely, as the scalar product of two complex *vectors* whose η th components are c_η' and c_η , respectively. Since c_η given by (2.2.18) is of the form of the projection of ϕ upon the ψ_η -axis, the wave function $\phi(q)$ can alternatively be represented by a vector $\{c_\eta\}$. Denoting this vector by a one *row* matrix $\langle\phi|$, or a one *column* matrix $|\phi\rangle$, one can rewrite (2.2.19) as

$$(\phi', \phi) = \langle\phi'| \phi\rangle, \quad (2.2.20)$$

namely, as the scalar product of “bra” vector $\langle\phi|$ and “ket” vector $|\phi\rangle$, thus closing the bracket: $\langle| \rangle$, as Dirac,¹ their “godfather”, joked. Corresponding to the

vector components $c_\eta \equiv (\psi_\eta, \phi)$, the *matrix elements* $A_{\eta'\eta}$ of an operator A are defined by

$$A_{\eta'\eta} \equiv (\psi_{\eta'}, A\psi_\eta) \equiv \langle \psi_{\eta'} | A | \psi_\eta \rangle \text{ or more briefly } \langle \eta' | A | \eta \rangle. \quad (2.2.21)$$

Thus, one can choose whichever *representation* one considers more convenient, the *operator acting on* the wave function, or the *matrix multiplied on* the vector.

The harmonic oscillator is the best and most tractable system to confirm this equivalence. With the use of (2.2.16), one obtains the matrix elements of the operator b^\dagger defined by (2.2.21) as

$$(b^\dagger)_{\eta'\eta} \equiv (\psi_{\eta'}, b^\dagger \psi_\eta) = \delta_{\eta', \eta+1} (\eta + 1)^{1/2}, \quad (2.2.22)$$

so that the matrices of b^\dagger and b can be written as

$$b^\dagger = \begin{pmatrix} 0 & \sqrt{1} & 0 & \dots \\ 0 & 0 & \sqrt{2} & \dots \\ 0 & 0 & 0 & \sqrt{3} \\ \dots & \dots & \dots & \dots \end{pmatrix} \quad b = \begin{pmatrix} 0 & 0 & 0 & \dots \\ \sqrt{1} & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & \dots \\ \dots & \dots & \sqrt{3} & \dots \end{pmatrix}. \quad (2.2.23)$$

The commutation relation (2.2.6) can now be confirmed by matrix algebra. The number operator $n = b^\dagger b$ has diagonal form:

$$n = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 2 & \dots \\ \dots & \dots & \dots & 3 \end{pmatrix} \quad (2.2.24)$$

as it should since the basis functions ψ_η defining the matrix elements are the eigenstates of n belonging to the respective eigenvalues η according to (2.2.9).

Switching back to the wave function, let us rewrite (2.2.14), with the use of (2.2.4), as a differential equation for $\psi_0(q)$:

$$-q/a^2 = (d\psi_0/dq)/\psi_0 = d(\ln \psi_0)/dq.$$

Integrating and normalizing, we obtain the ground-state wave function:

$$\psi_0(q) = (\pi a^2)^{-1/4} \exp[-q^2/2a^2]. \quad (2.2.25)$$

Application of eqs. (2.2.4, 2.2.17) gives the normalized excited-state wave functions:

$$\psi_\eta(q) = \{\eta! 2^\eta\}^{-1/2} \{q/a - a(d/dq)\}^\eta \psi_0(q) \quad (\eta = 1, 2, \dots). \quad (2.2.26)$$

Let us see how the expectation value of the position coordinate q varies with time, making use of eqs. (2.1.16), (2.2.4, 2.2.5), (2.2.18) and (2.2.22):

$$\begin{aligned}\langle q \rangle_t &\equiv (\psi(q, t), q \psi(q, t)) \\ &= 2^{-1/2} a \sum_{\eta} \sum_{\eta'} c_{\eta'}^*(0) c_{\eta}(0) \exp\{i(E_{\eta} - E_{\eta'})t/\hbar\} (\psi_{\eta}, \{b + b^{\dagger}\} \psi_{\eta'}) \\ &= \langle q \rangle_0 \cos \omega t + \langle p \rangle_0 \omega^{-1} \sin \omega t\end{aligned}\quad (2.2.27)$$

which agrees with the *classical* motion of a harmonic oscillator for the given initial conditions for q and p . Derivation of (2.2.27) in the Heisenberg picture is more straightforward since the Heisenberg equations of motion (2.1.10) for the Hamiltonian (2.2.1) gives the results:

$$dq/dt = (i/\hbar)[H, q]_- = p/m, \quad dp/dt = (i/\hbar)[H, p]_- = -m\omega^2 q, \quad (2.2.28)$$

which are exactly the same as the classical ones. Hence (2.2.27) follows immediately. Such an agreement between quantal and classical motion is characteristic of harmonic oscillation. In the Schrödinger picture, it is ascribed to the common energy difference $\hbar\omega$ between adjacent energy levels which are connected only by nonvanishing matrix elements of b, b^{\dagger} and hence of q , as is seen from the fact that only the products of the form $c_{\eta}^*(0)c_{\eta\pm 1}(0)$ contribute to $\langle q \rangle_0$ and $\langle p \rangle_0$. This quasi-classical feature, however, is absent in each eigenstate ψ_{η} in which $\langle q \rangle_t$ and $\langle p \rangle_t$ always vanish. In general, the classical picture of dynamics is found not in individual energy eigenstates – stationary states, but in the transitions or phase relations between them. This is the best example of the realization of the *correspondence principle* which had been advocated by Bohr and thereafter played an important guiding role towards the discovery of quantum mechanics (see Preface).

2.3 Uncertainty relation and complementarity

We found in Section 2.2 that the harmonic oscillator with Hamiltonian (2.2.1) has energy eigenvalues:

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega \quad (n = 0, 1, 2, \dots) \quad (2.3.1)$$

where we have rewritten the eigenvalue η of the number operator n defined by (2.2.7) simply as n (as is often done in applications of quantum mechanics), assuming that no confusion is caused thereby. Except for the additional term $\hbar\omega/2$ which is called the *zero-point energy*, this agrees with the working hypothesis (1.4.5) due to Planck, so that the temperature-dependent part of thermal radiation discussed there is exactly reproduced by quantum mechanics. The integral of the zero-point energy multiplied by the number density $8\pi\lambda^{-4}d\lambda$ of the normal modes κ , which diverges

as $\lambda \rightarrow 0$, is considered to belong to vacuum itself and hence is not observable. The observable part, $n\hbar\omega$, is interpreted as the ensemble of n photons with energy $\hbar\omega$ per mode, according to Einstein. Correspondingly, b and b^\dagger defined by (2.2.4) are called *annihilation* and *creation operators*, respectively, of a photon because of their matrix elements given by (2.2.22) or (2.2.23). Equation (2.3.1) indicates that any number of photons can be found per one mode (κ , see Section 1.3) or one state. Such particles are called *Bose particles* or, briefly, *bosons*. In the case of the harmonic oscillator of material particles, the zero-point energy is a quantity which is measurable: the dissociation energy of a molecule and the cohesive energy of a solid include a negative contribution from the zero-point vibrational energies of the atoms involved.

The zero-point energy of a harmonic oscillator is the result of the *uncertainty principle*. According to this general principle of quantum mechanics, any dynamical variables q and p which are canonically conjugate to each other, namely, which satisfy the commutation relation:

$$[q, p]_- \equiv qp - pq = i\hbar \quad (2.3.2)$$

(as is valid for q and p defined by (2.1.4)), can be determined simultaneously only with accuracies limited by the inequality, the so-called *uncertainty relation*:

$$\Delta q \cdot \Delta p \geq \hbar/2. \quad (2.3.3)$$

Here the uncertainty Δq in the state ψ is defined by a dispersion: $(\Delta q)^2 \equiv \langle (q - q_0)^2 \rangle_\psi$ where $q_0 \equiv \langle q \rangle_\psi$. The inequality (2.3.3) can be proved as follows. With a real number λ we define $\phi \equiv \{\lambda(q - q_0) + i(p - p_0)\}\psi$, and rewrite its scalar product with itself, making use of eqs. (2.1.7) and (2.3.2), as follows:

$$\begin{aligned} (\phi, \phi) &= (\psi, \{\lambda(q - q_0) - i(p - p_0)\}\{\lambda(q - q_0) + i(p - p_0)\}\psi) \\ &= (\Delta q)^2 \lambda^2 - \hbar \lambda + (\Delta p)^2 \\ &= \{\Delta q \lambda - \hbar/(2\Delta q)\}^2 - (\Delta q)^{-2} \{(\hbar^2/4 - (\Delta q)^2(\Delta p)^2)\}. \end{aligned} \quad (2.3.4)$$

Since the l.h.s. cannot be negative for any λ because of its definition, the second curly bracket $\{\cdot\cdot\}$ on the r.h.s. cannot be positive, whence follows the inequality (2.3.3).

This proof allows us also to see when the equality holds in (2.3.3). It is realized only when $\phi = 0$, namely

$$\{\lambda(q - q_0) + i(p - p_0)\}\psi = 0 \quad (2.3.5)$$

for the value $\lambda = \hbar/[2(\Delta q)^2]$ which makes the first curly bracket $\{\cdot\cdot\}$ on the r.h.s. of eq. (2.3.4) vanish. Equation (2.3.5) can be integrated with the use of (2.1.4). The normalized wave function is given by

$$\psi(q) = \{2\pi^{1/2}(\Delta q)^2\}^{-1/2} \exp\{-(q - q_0)^2/4(\Delta q)^2 - ip_0(q - q_0)/\hbar\}. \quad (2.3.6)$$

It reduces to the ground-state wave function of the harmonic oscillator, (2.2.25), when one puts $p_0 = q_0 = 0$ and $m^{-1/2} \Delta p = m^{1/2} \omega \Delta q = (\hbar \omega / 2)^{1/2}$. This indicates that the zero-point energy, $\hbar \omega / 2$, of the harmonic oscillator is *entirely* due to the uncertainty principle. Note that (2.3.6) includes also plane waves ($\Delta q \rightarrow \infty$) as well as completely localized state $\delta(q - q_0)$ ($\Delta q \rightarrow 0$). Ignoring mathematical rigor, one can consider the latter as an eigenfunction of the position operator q belonging to the eigenvalue q_0 : $q \delta(q - q_0) = q_0 \delta(q - q_0)$. What we called “ q -diagonal representation” below eq. (2.1.3) is that representation in which the operator q is diagonalized, namely the wave function is expanded on the basis of the eigenfunctions $\delta(q - q_0)$ with $(-\infty < q_0 < +\infty)$: $\psi(q) = \int dq_0 \psi(q_0) \delta(q - q_0)$.

Corresponding to (2.3.3), there is another uncertainty relation between time and energy:

$$\Delta t \Delta E \geq \hbar / 2. \quad (2.3.7)$$

One should note an important difference between (2.3.3) and (2.3.7): the position coordinate q is a q -number (operator) while the time coordinate t remains a c -number in quantum mechanics. In fact, the classical uncertainty relation between time and angular frequency:

$$\Delta t \Delta \omega \geq 1/2 \quad (2.3.8)$$

was well known long before Planck’s discovery of h and played an important role in spectroscopy. Putting the fourth component of the Einstein–de Broglie relation $(\mathbf{p}, E) \rightleftharpoons (\hbar \mathbf{k}, \hbar \omega)$ between particle and wave (which played a pivotal role in the discovery of quantum mechanics) into (2.3.8), one readily obtains (2.3.7).

The classical uncertainty relation, (2.3.8), is obtained as follows. By superposing oscillations with different angular frequencies ω by a suitable ω -dependent amplitude $f(\omega)$, one can form a pulse $g(t)$ in the time domain:

$$g(t) = \int_{-\infty}^{+\infty} \exp(-it\omega) f(\omega) d\omega. \quad (2.3.9)$$

As a simple example, let us take a Gaussian spectral distribution

$$f(\omega) \propto \exp[-(\omega - \omega_0)^2 / 4(\Delta \omega)^2]$$

characterized by dispersion $\Delta \omega$ in its intensity $|f(\omega)|^2$:

$$(\Delta \omega)^2 \equiv \int_{-\infty}^{+\infty} (\omega - \omega_0)^2 |f(\omega)|^2 d\omega / \int_{-\infty}^{+\infty} |f(\omega)|^2 d\omega. \quad (2.3.10)$$

The Fourier transform (2.3.9) gives an intensity $|g(t)|^2 \propto \exp[-2(\Delta \omega)^2 t^2]$ with dispersion $\Delta t = (2\Delta \omega)^{-1}$ according to the definition (2.3.10), which corresponds to the minimum possible value of $\Delta t \Delta \omega$ in (2.3.8). The inequality (2.3.8) for

a general form of $f(\omega)$ can be derived from the Fourier transform (2.3.9), or can readily be imagined in analogy with the derivation of (2.3.6) as the case of minimum uncertainty between q and p .

A commutation relation similar to (2.1.3) holds between the phase ϕ (to be defined later) of an electromagnetic wave and the number n of photons of the same mode:

$$[\phi, n]_- = i, \quad (2.3.11)$$

which is equivalent to putting $n = -i\partial/\partial\phi$ (compare with (2.1.4)). Applying (2.3.11) or its equivalent to the power series expansion of $\exp(-i\phi)$, one obtains

$$[\exp(-i\phi), n]_- = \exp(-i\phi) = i\partial/\partial\phi\{\exp(-i\phi)\}. \quad (2.3.12)$$

Let us now define the phase ϕ by

$$b = \exp(-i\phi)n^{1/2}. \quad (2.3.13)$$

This definition is partly legitimized by the fact that the scalar product of the conjugate operator: $b^\dagger = n^{1/2} \exp(i\phi)$ with the operator (2.3.13) (paying attention to this order of multiplication) gives immediately $b^\dagger b = n$ in agreement with the definition of n , eq. (2.2.7). The commutation relation between b and n :

$$[b, n]_- = b \quad (2.3.14)$$

which follows from (2.2.6), turns out to be equivalent to (2.3.12). In fact, multiplying $n^{1/2}$ from the right to (2.3.12) and using the definition (2.3.13) one obtains (2.3.14). Thus we have a posteriori confirmed (2.3.12) and hence (2.3.11).

The commutation relation (2.3.11) leads to the uncertainty relation

$$\Delta\phi\Delta n \geq \frac{1}{2} \quad (2.3.15)$$

in the same way that (2.3.2) led to (2.3.3). Since the wave picture and the corpuscular picture are valid for $\Delta\phi \ll 1$ and $\Delta n \ll 1$, respectively, the inequality (2.3.15) means that the two pictures for the electromagnetic field are not simultaneously valid but rather are complementary to each other. This *complementarity* is the quantum mechanics answer to the “wave–particle duality” which was controversial throughout the age of the old quantum theory. **N.B.** The word “old quantum theory” is popularly used among physicists, being distinguishable from quantum mechanics. Old quantum theory consists of various provisional theories and models framed between 1900 and 1925.

2.4 Density matrix

In this section, we will give a mathematical framework with which we can describe quantum-mechanical states in more general situations. Readers new to this field are advised to make use of this section only as a reference since the use of the density matrix introduced here will be avoided in this book (except in a few sections) for pedagogical reasons. After having read the main part of the book they may find this section to be useful and possibly of profound interest.

Let us begin by introducing some kinds of quantum-mechanical operators and the theorems they satisfy.

If an Hermite operator A satisfies the inequality

$$(\psi, A\psi) > 0 \quad \text{or} \quad (\psi, A\psi) \geq 0 \quad (2.4.1)$$

for any function $\psi(q)$, A is called a *positive-definite* or *non-negative* operator. The Hamiltonian (2.2.1) for a harmonic oscillator is such an example. Because of the Hermiticity of p and q , we have

$$(\psi, p^2\psi) = (p\psi, p\psi) \geq 0, \quad (\psi, q^2\psi) = (q\psi, q\psi) \geq 0$$

and hence $(\psi, H\psi) \geq 0$, non-negative. (In fact, H is positive-definite because of zero-point vibration.)

Let us now define the *trace* of an Hermite operator A by the sum of all its eigenvalues a_n :

$$\text{Tr}(A) \equiv \sum_n a_n = \sum_n (\psi_n, A\psi_n) \quad (2.4.2)$$

where $\psi_n(q)$ is the eigenstate corresponding to a_n . Consider another orthonormal complete set of functions $\{u_k(q)\}$ such as, for example, the plane-wave eigenstates of the free particle Hamiltonian $K = p^2/2m$ given in Section 1. 3, satisfying in general the relations

$$(u_k, u_{k'}) = \delta_{kk'} \quad (2.4.3)$$

which is of the same form as

$$(\psi_n, \psi_{n'}) = \delta_{nn'}. \quad (2.4.4)$$

One can expand ψ_n in terms of $\{u_k\}$ as

$$\psi_n(q) = \sum_k c_{nk} u_k(q). \quad (2.4.5)$$

Inserting (2.4.5) into (2.4.4) and making use of (2.4.3), one obtains the orthonormality condition for the expansion coefficients:

$$\sum_k c_{nk}^* c_{n'k} = \delta_{nn'}. \quad (2.4.6)$$

Putting (2.4.5) into (2.4.2) and making use of (2.4.6), one obtains

$$\text{Tr}(A) = \sum_k \sum_{k'} \sum_n c_{nk}^* c_{n'k'} (u_k, Au_{k'}) = \sum_k (u_k, Au_k). \quad (2.4.7)$$

Comparison of the r.h.s. of (2.4.2) and (2.4.7) indicates that $\text{Tr}(A)$ is the sum of the *diagonal elements* of A , being *independent of (invariant against)* the choice of the orthonormal set with which to take the matrix element. This indicates the physical significance of $\text{Tr}(A)$ as is exemplified by the partition function Z and the free energy F mentioned below, and one can calculate it without necessarily solving the eigenvalue problem of A .

We will give here some theorems concerning the trace of the products of operators. In the first place, we have

$$\text{Tr}(AB) = \sum_k (AB)_{kk} = \sum_k \sum_{k'} A_{kk'} B_{k'k} = \sum_{k'} (BA)_{k'k'} = \text{Tr}(BA). \quad (2.4.8)$$

By repeated applications of this rule to the triple product, we obtain

$$\text{Tr}(ABC) = \text{Tr}[A(BC)] = \text{Tr}[(BC)A] = \text{Tr}(BCA) = \text{Tr}(CAB). \quad (2.4.9)$$

Namely, the order of operators in their product can be changed only *cyclically*, not arbitrarily, within the trace operation.

With the harmonic oscillator, one can confirm (2.4.1) for each eigenfunction ψ_n of H since $(\psi_n, H\psi_n) = a_n = (n + \frac{1}{2})\hbar\omega > 0$. However, the summation in (2.4.2) diverges to infinity. As an example with a finite value of trace, let us consider an exponential operator $A \equiv \exp(-\beta H)$ with positive β . The operator A is so defined that it has eigenvalues $\exp(-\beta E_n)$ with eigenfunctions $\{\psi_n\}$ being the same as those of H . With H of the harmonic oscillator, we have in fact

$$\begin{aligned} Z &\equiv \text{Tr}[\exp(-\beta H)] \quad (\text{definition for general } H) \\ &= \exp(-\beta\hbar\omega/2)[1 - \exp(-\beta\hbar\omega)]^{-1} < \infty. \\ &\quad (\text{for a harmonic oscillator}). \end{aligned} \quad (2.4.10)$$

If β represents the reciprocal temperature defined in (1.4.1), the normalized positive definite operator defined by

$$\rho \equiv \exp(-\beta H)/\text{Tr}[\exp(-\beta H)], \quad \text{with } \text{Tr}(\rho) = 1 \quad (2.4.11)$$

represents the quantal version of the classical Boltzmann distribution function given by (1.4.1) in the sense that the thermodynamical *average* of the quantum-mechanical

expectation values of an arbitrary physical quantity B is given by

$$\begin{aligned}\langle\langle B \rangle\rangle &= \sum_n \rho_n B_n = \sum_n \sum_m \rho_n \delta_{nm} B_{mn} \\ &= \sum_n (\rho B)_{nn} = \text{Tr}(\rho B).\end{aligned}\quad (2.4.12)$$

Hamiltonians of all realistic physical systems are considered to have finite Z since the number density dn/dE_n of eigenstates usually varies with a power function of E_n asymptotically with increasing E_n . It can be shown that the free energy F of the system is given by

$$F(T) = -\beta^{-1} \ln Z(\beta). \quad (2.4.13)$$

Let us now study how the expectation value of a physical quantity B , given by (2.4.12), changes with time t under the time-dependent Hamiltonian H_t . According to eq. (2.1.10) representing the Heisenberg picture, B_t is subject, within a small time interval Δt , to a small change given by

$$i\hbar \Delta B_t = [B_t, H_t]_- \Delta t. \quad (2.4.14)$$

Since the state does not change according to this picture, we have a small change of the expectation value (statistically averaged) given by

$$i\hbar \Delta \langle\langle B_t \rangle\rangle = i\hbar \text{Tr}(\rho \Delta B_t) = \text{Tr}(\rho [B_t, H_t]_- \Delta t). \quad (2.4.15)$$

With the use of the cyclic exchange rule (2.4.9), the r.h.s. of (2.4.15) can be rewritten as

$$\text{Tr}(\{[H_t, \rho]_- \Delta t\} B_t).$$

According to the Schrödinger picture in which one ascribes the change in the expectation value to the change in the state $\Delta\rho$, the l.h.s. of (2.4.15) can be written as

$$i\hbar \text{Tr}(\{\Delta\rho\} B_t).$$

In order that the change in the expectation value of the physical quantity is independent of the picture used, one must equate this change in the Schrödinger picture to (2.4.15) in the Heisenberg picture. Dividing both sides by Δt , one obtains a differential equation

$$i\hbar d\rho_t/dt = [H_t, \rho_t]_- \quad (2.4.16)$$

in the Schrödinger picture, where the suffix t to ρ is revived since the latter is now a function of time. As in the case of the wave function description, the time derivative of ρ_t in the Schrödinger picture is opposite to that of the physical quantity B_t in the Heisenberg picture.

Let us start with the system in thermal equilibrium as given by (2.4.11). Because H_t is different from H which governed the thermal equilibrium, the system is brought into a non-equilibrium state. The density matrix ρ_t representing this state can be quite general except that it is a normalized positive-definite operator. Therefore, the most general state of a system is represented by a normalized non-negative density matrix, not by a wave function. What does the wave function mean from this generalized viewpoint? To see this, let us forget the time dependence for the moment, and note that *any* normalized non-negative Hermite operator ρ has non-negative eigenvalues w_n and eigenstates $\psi_n(q)$ satisfying the eigenvalue equations

$$\rho\psi_n(q) = w_n\psi_n(q), \quad (2.4.17)$$

and the normalization relation

$$\sum_n w_n = 1. \quad (2.4.18)$$

For example, the eigenfunctions of the density matrix (2.4.10) for thermal equilibrium are those of the Hamiltonian H . In the q -diagonal representation (see Section 2.1), a general density matrix can be written as

$$\rho(q, q') = \sum_n w_n \psi_n(q) \psi_n^*(q'), \quad (2.4.19a)$$

an integral kernel with which (2.4.17) can be rewritten as

$$\int \rho(q, q') \psi_n(q') dq' = w_n \psi_n(q). \quad (2.4.17a)$$

Alternatively, with the use of ket and bra vectors, ρ can be written as

$$\rho = \sum_n |\psi_n\rangle w_n \langle\psi_n| \quad (2.4.19b)$$

with which (2.4.17) can be written as

$$\rho|\psi_n\rangle = w_n|\psi_n\rangle. \quad (2.4.17b)$$

Let us now consider an operator equation

$$\rho(1 - \rho) = 0. \quad (2.4.20)$$

Since the operator on the l.h.s. has the same set of eigenfunctions as ρ itself, (2.4.20) is equivalent to

$$w_n(1 - w_n) = 0 \quad (\text{for all } n). \quad (2.4.21)$$

Because of the normalization relation (2.4.18), eq. (2.4.21) means that $w_n = 1$ for only one n , say 1, and zero for all other n s. Such a state ρ_1 , which is called a “pure state” in contradistinction from a “mixed state” expressed by (2.4.19a, b),

is equivalent to what we expressed by the wave function $\psi_1(q)$. Thus, the wave function can represent only a restricted class among all possible states of a system.

Let us come back to our time-dependent problem. The evolution of the density matrix (2.4.19a) according to its equation of motion (2.4.16) is such that each $\psi_n(q, t)$ follows the time-dependent Schrödinger equation

$$i\hbar d\psi_n(q, t)/dt = H_t \psi_n(q, t) \quad (2.4.22)$$

while w_n is kept constant. The latter fact has often been erroneously used to argue against the irreversibility of a system. Putting aside such a profound problem of statistical mechanics, let us ask what physical meaning this time-dependent $\psi_n(q, t)$ has. It is not an instantaneous eigenstate of H_t when the latter varies with time. What has been done in most applications of quantum mechanics to time-dependent problems is not a consecutive pursuit of a single wave function but that of its expansion coefficients as projected on a fixed (time-independent) complete set of orthonormal wave functions. The eigenfunctions of important physical quantities such as energy, momentum and angular momentum are especially useful for such purposes, as will be shown by many examples in this book. Sometimes, it is more useful to pursue the evolution of the density matrix itself instead of that of the wave function (see Sections 11.3–11.5). Even in such a case, the calculations can be performed only with the use of a fixed set of orthonormal functions.

Consider two interacting systems, I and II. It can be shown by a simple argument that the state of the subsystem (I or II) is in a mixed state even when the composite system (I + II) is in a pure state. An obvious example is the matter system (I) in thermal equilibrium with a heat bath (II), which is in fact in the mixed state (2.4.11). This fact is of particular importance when we describe the measurement of a physical quantity of a microscopic system (I) by a measuring apparatus (II). Well-known paradoxes concerning quantum-mechanical descriptions of a measuring process can be shown⁴ to originate from an unjustifiable description of the microscopic system in terms of a pure state.

It must be mentioned, in this context, that the statistical distribution of the results of measuring a microscopic quantity inherent in quantum mechanics on the one hand, and the statistical distribution of the state of a matter system in contact with a heat bath on the other, share the same origin. Although the quantum-mechanical expectation value and the thermodynamical average are shown respectively by $\langle \dots \rangle$ and $\langle\langle \dots \rangle\rangle$ in this book to indicate different procedures of calculation, they have the same origin when everything (including the heat bath as well as the measuring apparatus) is described by quantum mechanics.

3

Interaction of radiation with matter

3.1 *A* and *B* coefficients of Einstein

Absorption and emission of a photon by matter are the simplest elementary processes of radiation–matter interaction. *Before* the establishment of quantum mechanics, Einstein¹ put forth a general argument concerning them with the use of a model consisting of a photon field and many identical atoms all with two energy levels: a ground state E_1 and an excited state E_2 , contained in a cavity. The relationships he derived for the rates of absorption, spontaneous emission and induced emission have been confirmed by quantum mechanics and are still valid.

He assumed that absorption and emission of photons take place with energy $\hbar\omega = E_2 - E_1$ by excitation and de-excitation of the atoms, with the rates (transition probabilities per unit time) $N_1 B_{12}w(\omega)$ and $N_2 B_{21}w(\omega)$, respectively, where N_1 and N_2 denote the number densities of atoms in the ground and excited states, respectively, and $w(\omega)$ the spectral energy density of radiation at ω which will be specified later. He further considered the *spontaneous* emission of a photon by excited-state atoms with the rate $N_2 A_{21}$, which is *independent* of $w(\omega)$ by definition. This term *must* be considered in order for thermal equilibrium between radiation and matter to be realized, as will be seen. The rate equation is then given by

$$dN_1/dt = -dN_2/dt = N_2 A_{21} - N_1 B_{12}w(\omega) + N_2 B_{21}w(\omega). \quad (3.1.1)$$

Einstein considered the general situation where $w(\omega)$ consists of an *incident* electromagnetic field and a *thermal* radiation field.

In the absence of an incident field, (3.1.1) should have a solution in which atoms are in thermal equilibrium with the thermal radiation field with temperature T . Putting $dN_1/dt = 0$, one obtains

$$w(\omega) = A_{21}/[(N_1/N_2) B_{12} - B_{21}]. \quad (3.1.2)$$

In thermal equilibrium, the population of atoms obeys the Boltzmann distribution:

$$N_2/N_1 = g_2 \exp(-E_2/k_B T)/g_1 \exp(-E_1/k_B T) = (g_2/g_1) \exp(-\hbar\omega/k_B T) \quad (3.1.3)$$

where g_1 and g_2 are the degeneracies of the states 1 and 2. In order that (3.1.2) agrees with, W_T , the spectral energy density (1.4.9) of the thermal radiation field, the following relations must be satisfied:

$$g_1 B_{12} = g_2 B_{21}, \quad (\hbar\omega^3/\pi^2 c_0^3) B_{21} = A_{21}. \quad (3.1.4)$$

The coefficients A_{21} , B_{12} , B_{21} for the three processes are called the Einstein coefficients. The relations (3.1.4) among them hold quite generally, irrespective of the origin of the radiation field concerned. Quantum-mechanical calculation of the individual coefficients for typical radiation-matter interactions will be presented in a later section, which in fact satisfy the Einstein relations (3.1.4).

Let us compare the rates of two kinds of emission from an excited-state atom, spontaneous emission, and induced emission due to a thermal radiation field.² They are given by A_{21} and $B_{21}w(\omega) = A_{21}n(\omega)$ (see eqs. (1.4.9) and (3.1.4)), respectively. Their ratio, $n(\omega)^{-1} = \exp(\hbar\omega/k_B T) - 1$, is greater or smaller than unity according as $\hbar\omega >$ or $< 0.693 k_B T \equiv \hbar\omega_T$. Since the corresponding wavelength λ_T is 70 μm at room temperature, spontaneous emission is predominant in the infrared and shorter wavelength region whereas thermally induced emission dominates in the region where the wavelength is longer than the far infrared.

If the incident field with a mode characterized by a definite wave vector \mathbf{k} is much greater than the thermal radiation field, the third term on the r.h.s. of (3.1.1) represents the induced emission of radiation in the same direction as the incident wave. If the population is kept inverted ($N_2/g_2 > N_1/g_1$) by external pumping in a cavity in which the radiation field is confined, this term overcomes the second term, resulting in an *indefinite amplification* of a few particularly favorable modes of the electromagnetic wave due to self-organization. This is the principle of the *laser*, the intense and coherent light source which plays an important role in modern spectroscopy and in a variety of technological applications. The increasing energy of the electromagnetic waves is supplied by an external source of pumping through atoms excited thereby as they are de-excited by the overwhelming induced emission, while the initial seed of the amplification, before the \mathbf{k} wave is established, is the spontaneous emission of pumped atoms.

3.2 Relation between absorption and emission spectra in thermal equilibrium

In molecules and insulating solids, one often encounters the situation in which optical transitions take place between a group of ground states $\{i\}$ and a group of excited states $\{j\}$ which energetically are well separated from each other but are close by within each group. In such a situation, the absorption (as well as the emission) spectrum consists of a band of densely distributed lines with $\hbar\omega_{ji} = E_j - E_i$. As a typical example, let us consider a molecule. A group represents an electronic state, the ground or the excited state which are typically separated by orders of eV, and members within a group consist of vibrational states which are typically separated by the order of 0.01 eV since the atoms are much heavier than the electrons. With each transition ($i \rightleftharpoons j$) are associated Einstein's A and B coefficients which satisfy the relations (3.1.4).

Assuming a Boltzmann distribution in the ground states: $p_i = Z^{-1}g_i \times \exp(-\beta E_i)$, $Z \equiv \sum_i g_i \exp(-\beta E_i)$ and neglecting the population in the excited states because the excitation energies $\hbar\omega_{ji}$ are much larger than $k_B T$, the absorption spectrum is given by

$$B(E) = \sum_i \sum_j Z^{-1} g_i \exp(-\beta E_i) B_{ij} \delta(E_j - E_i - E). \quad (3.2.1)$$

When matter is excited into any of the excited states comprising group $\{j\}$, it will be promptly (typically in 10^{-12} s) repopulated towards the Boltzmann distribution *within* that group $\{j\}$ of excited states, since the spontaneous emission therefrom takes a much longer time (10^{-8} s, see eq. (3.5.7)). Then the luminescence spectrum is given by

$$A(E) = \sum_i \sum_j Z'^{-1} g_j \exp(-\beta E_j) A_{ji} \delta(E_j - E_i - E) \quad (3.2.2)$$

where $Z' \equiv \sum_j g_j \exp(-\beta E_j)$.

With the use of the relations (3.1.4), one finds that the emission and the absorption spectra are related by

$$A(E) = (E^3 / \pi^2 \hbar^2 c^3) (Z' / Z) \exp(-\beta E) B(E) \quad (3.2.3)$$

where the light velocity c_0 in a vacuum is replaced by $c = c_0 / n(\omega)$ in a medium with index of refraction $n(\omega)$ at that frequency ω . Note that we are using the same notation $n(\omega)$ for the refractive index and the number of quanta (photons and phonons) in accordance with historical convention, hoping no confusion is caused thereby.

The relation (3.2.3) will prove to be useful not only for the comparative study of the two spectra, but also as a criterion to confirm to what extent thermalization among the excited states has been achieved during their radiative lifetime after

optical excitation. Under the assumption of complete thermalization, (3.2.3) indicates that at absolute zero of temperature the absorption and emission spectra appear only on the high and low energy sides, respectively, of a *common energy point* E_0 where a line spectrum (such as a zero-phonon line) can appear.

A typical system for which the relation (3.2.3) can be realized is a localized electron in a solid as is described in Section 4.6. However, if the photo-excited electron has insufficient time to be thermalized before it emits a photon, as in the case of resonant second-order optical processes to be described in Section 11.4, the deviation from (3.2.3) appears as a *hot luminescence* component, as is most clearly seen in the figures in Appendix 2. The polariton bottleneck of the exciton luminescence to be described in Section 10.8 is an *intrinsic* situation in which thermalization can never be realized.

3.3 Canonical form of radiation–matter interaction

It is convenient to make use of the canonical formalism for the radiation–matter interaction to describe the optical processes in quantum mechanics.³ Consider a particle with charge e with position \mathbf{r} and velocity \mathbf{v} in an electromagnetic field, which is subject to the Lorentz force:

$$\mathbf{F} = e[\mathbf{E}(\mathbf{r}, t) + \mathbf{v} \times \mathbf{B}(\mathbf{r}, t)] \quad (3.3.1)$$

as is obtained from (1.1.7) under the assumption that the charge density $\rho(\mathbf{r})$ is concentrated at the particle position \mathbf{r} . The classical-mechanical motion of this particle with mass m can be described in terms of the Hamiltonian:

$$H = (\mathbf{p} - e\mathbf{A}(\mathbf{r}, t))^2/2m + e\phi(\mathbf{r}, t) \quad (3.3.2)$$

where $\mathbf{A}(\mathbf{r}, t)$ and $\phi(\mathbf{r}, t)$ are respectively vector and scalar potentials related to $\mathbf{B}(\mathbf{r}, t)$ and $\mathbf{E}(\mathbf{r}, t)$ through (1.3.1) and (1.3.3). To confirm that (3.3.2) is the appropriate Hamiltonian, let us write down the canonical equations of motion derived therefrom:

$$\begin{aligned} d\mathbf{r}/dt &= \partial H/\partial \mathbf{p} = [\mathbf{p} - e\mathbf{A}(\mathbf{r}, t)]/m, \\ dp_x/dt &= -\partial H/\partial x = (e/m)[\mathbf{p} - e\mathbf{A}(\mathbf{r}, t)] \cdot \partial \mathbf{A}/\partial x - e\partial\phi/\partial x, \text{ etc.} \end{aligned} \quad (3.3.3)$$

Inserting (3.3.3) and its time-derivative into both sides of the latter equation, one obtains

$$\begin{aligned} md^2x/dt^2 &= e[(-\partial A_x/\partial t - \partial\phi/\partial x) + dy/dt(\partial A_y/\partial x - \partial A_x/\partial y) \\ &\quad - dz/dt(\partial A_x/\partial z - \partial A_z/\partial x)] = e[E_x + (dy/dt)B_z - (dz/dt)B_y]. \end{aligned}$$

This is nothing other than the Newtonian equation of motion for the charged particle since the r.h.s. is the Lorentz force (3.3.1).

The quantum-mechanical Hamiltonian for the system composed of interacting charged particles (e_i, m_i, \mathbf{r}_i) and electromagnetic waves can then be written as

$$H = \sum_i [\mathbf{p}_i - e_i \mathbf{A}(\mathbf{r}_i)]^2 / 2m_i + \sum \sum_{i < j} e_i e_j / [4\pi\epsilon |\mathbf{r}_i - \mathbf{r}_j|] + \sum_\lambda n_\lambda \hbar \omega_\lambda$$

$$= H_M + H_R + H_I^{(1)} + H_I^{(2)} \quad (3.3.4)$$

where

$$H_M = \sum_i \mathbf{p}_i^2 / 2m_i + \sum \sum_{i < j} e_i e_j / [4\pi\epsilon |\mathbf{r}_i - \mathbf{r}_j|], \quad (3.3.5)$$

$$H_R = \sum_\lambda n_\lambda \hbar \omega_\lambda, \quad (3.3.6)$$

$$H_I^{(1)} = - \sum_i (e_i / m_i) \mathbf{p}_i \cdot \mathbf{A}(\mathbf{r}_i)$$

$$= - \sum_i \sum_\lambda (\hbar / 2\epsilon \omega_\lambda V)^{1/2} (e_i / m_i) (e_\lambda \cdot \mathbf{p}_i)$$

$$\times [b_\lambda \exp(i\mathbf{k} \cdot \mathbf{r}_i) + b_\lambda^\dagger \exp(-i\mathbf{k} \cdot \mathbf{r}_i)], \quad (3.3.7)$$

$$H_I^{(2)} = \sum_i (e_i^2 / 2m_i) \mathbf{A}(\mathbf{r}_i)^2. \quad (3.3.8)$$

In the above we have made use of eqs. (1.3.8, 1.3.10, 1.3.13) and (2.2.3, 2.2.5, 2.2.8) (the zero-point energy is removed for the reason mentioned in Section 2.3) and of the commutation relation:

$$\mathbf{p}_i \cdot \mathbf{A}(\mathbf{r}_i) - \mathbf{A}(\mathbf{r}_i) \cdot \mathbf{p}_i = -i\hbar \nabla_i \cdot \mathbf{A}(\mathbf{r}_i) = 0 \quad (3.3.9)$$

which is valid because of the transverse nature of the electromagnetic waves. It may be more convenient to rewrite (3.3.7) in the integral form with the use of the current density operator $\mathbf{j}(\mathbf{r})$:

$$H_I^{(1)} = - \int d\mathbf{r} \mathbf{j}(\mathbf{r}) \cdot \mathbf{A}(\mathbf{r})$$

$$= - \sum_\lambda (\hbar / 2\epsilon \omega_\lambda V)^{1/2} \mathbf{e}_\lambda \cdot (\mathbf{j}_\lambda b_\lambda + \mathbf{j}_\lambda^\dagger b_\lambda^\dagger) \quad (3.3.7a)$$

$$\text{where } \mathbf{j}(\mathbf{r}) \equiv \sum_i (e_i / m_i) (1/2) [\delta(\mathbf{r} - \mathbf{r}_i) \mathbf{p}_i + \mathbf{p}_i \delta(\mathbf{r} - \mathbf{r}_i)], \quad (3.3.10)$$

$$\mathbf{j}_\lambda \equiv \int \mathbf{j}(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}) d\mathbf{r}. \quad (3.3.11)$$

The symmetrization procedure for the non-commuting \mathbf{r}_i and \mathbf{p}_i as given by (1/2) $[\dots]$ in (3.3.10) is generally needed to assure the Hermiticity of the operator $\mathbf{j}(\mathbf{r})$ although it is immaterial in the present case of a transverse field (see (3.3.9)).

Equation (3.3.7) contains annihilation and creation operators b_λ and b_λ^\dagger of a photon (see the statement after eq. (2.3.1)), contributing to the absorption and emission of light in its first-order perturbation, as will be shown in Section 3.5.

Equation (3.3.8), which is quadratic in b_λ and b_λ^\dagger , contributes, in its first-order perturbation, to two-photon processes such as light scattering, together with the second-order perturbation of $H_I^{(1)}$.

3.4 First-order optical processes: absorption and emission of photon

Let us denote the eigenenergy and eigenstate of the matter by E_a and $|a\rangle$, respectively, and those of the radiation field by $E(\dots, n_\lambda, \dots)$ and $|\dots, n_\lambda, \dots\rangle$, respectively. One can then write

$$H_M|a\rangle = E_a|a\rangle, \quad (3.4.1)$$

$$H_R|\dots, n_\lambda, \dots\rangle = E(\dots, n_\lambda, \dots)|\dots, n_\lambda, \dots\rangle, \quad (3.4.2)$$

$$E(\dots, n_\lambda, \dots) = \sum_\lambda n_\lambda \hbar \omega_\lambda. \quad (3.4.3)$$

Here n_λ denotes an eigenvalue of the operator n_λ in (3.3.6). The same letter is used, hoping that will not cause any confusion.

One can now calculate the rate of transition of the non-perturbed system: $H_0 \equiv H_M + H_R$ from its initial state $|i\rangle = |a\rangle \dots, n_\lambda, \dots\rangle$ to either of the final states $|f\rangle = |b\rangle |\dots, n_\lambda \mp 1, \dots\rangle$, corresponding to absorption or emission of a photon of mode λ specified by wave vector \mathbf{k} and polarization vector \mathbf{e}_λ , through the perturbation $H_I^{(1)}$, with the use of the well-known formula,⁴ the so-called Fermi's golden rule (or also as will be derived in Section 3.5 in the context of a comparison between classical and quantal derivations):

$$\begin{aligned} W_{\text{abs. (or } W_{\text{em.}})} &= (2\pi/\hbar) \sum_f |(H_I^{(1)})_{fi}|^2 \delta[\{E_b + E(\dots, n_\lambda \mp 1, \dots)\} \\ &\quad - \{E_a + E(\dots, n_\lambda, \dots)\}] \\ &= \sum_\lambda \delta(\omega_{ba} \mp \omega_\lambda) \{n_\lambda \text{ or } (n_\lambda + 1)\} \\ &\quad \times \{\pi/(\hbar \epsilon \omega_\lambda V)\} |\mathbf{e}_\lambda \cdot (\mathbf{j}_{\pm \mathbf{k}})_{ba}|^2 \end{aligned} \quad (3.4.4)$$

where we have defined: $\hbar \omega_{ba} \equiv E_b - E_a$, and assigned the upper and lower of the \mp and \pm signs to the absorption and emission processes, respectively. If we take $E_b > E_a$, a and b in (3.4.4) as well as in the definitions of $|i\rangle$ and $|f\rangle$ should be replaced by b and a , respectively for the emission process. Then the expression $|\dots|$ in (3.4.4) is identical for either process.

Let us consider an incident photon flux within a small solid angle $\Delta\Omega$ around a given direction of propagation \mathbf{k}/k , polarization \mathbf{e}_λ and having a spectral energy density given by

$$w^{(\mathbf{e}, \mathbf{k})}(\omega) = n_\lambda \hbar \omega_\lambda \rho(\omega) \Delta\Omega. \quad (3.4.5)$$

Then the rate of absorption of a photon is given, according to (3.4.4), by

$$W_{\text{abs.}}^{(e,k)} = B_{ab}^{(e,k)} w^{(e,k)}(\omega_{ba}) \quad (3.4.6)$$

where

$$B_{ab}^{(e,k)} = (\pi/\epsilon)(\hbar\omega_{ba})^{-2} |e_\lambda \cdot (j_k)_{ba}|^2. \quad (3.4.7)$$

That part of the emission rate (3.4.4) which is proportional to n_λ represents the induced emission and is denoted by $W_{\text{ind. em.}}^{(e,k)}$. It is exactly equal to eq. (3.4.6). That part of (3.4.4) which is proportional to unity represents the spontaneous emission since it arises independently from (and hence even in the absence of) incoming electromagnetic waves. A photon can be emitted in any direction \mathbf{k}/k so long as the energy conservations: $\omega(=ck) = \omega_{ba}$ is satisfied. Its polarization, however, cannot be perpendicular to the vector: $(\dots)_{ba}$ of (3.4.7).

To see this point more clearly, let us neglect \mathbf{k} in the exponent of (3.4.7) by assuming that the corresponding wavelength is much greater than the extension of the electronic or atomic wave functions represented by $(\mathbf{r}_i)_{ba}$. Writing down the Heisenberg equation of motion for the system M:

$$p_{ix}/m_i \equiv dx_i/dt = i\hbar^{-1}[H_M, x_i]_- \quad (3.4.8)$$

and taking its matrix element, one obtains

$$(p_{ix})_{ba}/m_i = (i/\hbar)\{E_b(x_i)_{ba} - (x_i)_{ba}E_a\} = i\omega_{ba}(x_i)_{ba}. \quad (3.4.9)$$

In the case of *neutral* atoms and molecules, one can define the electric dipole moment

$$\boldsymbol{\mu} = \sum_i e_i \mathbf{r}_i \quad (3.4.10)$$

as an operator independent of the choice of the origin of the position coordinate since $\sum e_i = 0$. Noting that the atomic or molecular size is much smaller than the wavelength $\lambda = 2\pi/k$ of the electromagnetic wave except in X-ray and high-energy regions, one can approximate the matrix element in (3.4.7) as

$$(j_k)_{ba} \simeq i\omega_{ba}\boldsymbol{\mu}_{ba} \quad (3.4.11)$$

where $\boldsymbol{\mu}_{ba}$ is called the transition dipole moment. This approximation (neglect of \mathbf{k}) is therefore called the *dipole approximation*. Note that the non-diagonal matrix element $\boldsymbol{\mu}_{ba}$ can be defined also for ionized atoms and molecules since it is independent of the origin of position.

While the dipole moment is inherent in the material system M with its eigenstates a and b , the energy flux of radiation spontaneously emitted in the direction \mathbf{k}/k and with polarization e_λ is proportional to $(e_\lambda \cdot \boldsymbol{\mu}_{ba})^2$. Since e_λ is perpendicular to \mathbf{k} , the flux varies as $\sin^2 \theta$ where θ is the angle between \mathbf{k} and $\boldsymbol{\mu}_{ba}$, and the polarization

should be in the plane spanned by $\boldsymbol{\mu}_{ba}$ and \mathbf{k} . Taking the directional average: $\langle \sin^2 \theta \rangle = 2/3$ and making use of (1.3.7) without the factor 2 for polarization, one obtains

$$A_{ba} \equiv W_{\text{spont. em.}} = (\omega_{ba}^3 / 3\pi\epsilon \hbar c^3) |\boldsymbol{\mu}_{ba}|^2. \quad (3.4.12)$$

If M consists of identical molecules randomly oriented, we can take the directional average: $\langle |\mathbf{e}_\lambda \cdot \boldsymbol{\mu}_{ba}|^2 \rangle = |\boldsymbol{\mu}_{ba}|^2 / 3$. The rate of absorption and induced emission is then given by

$$W_{\text{abs.}} = W_{\text{ind. em.}} = B_{ab} w(\omega_{ba}), \quad (3.4.13)$$

$$B_{ab} = (\pi / 3\epsilon \hbar^2) |\boldsymbol{\mu}_{ba}|^2. \quad (3.4.14)$$

Thus the relation (3.1.4) is confirmed by the quantum-mechanical calculation which takes explicit account of directional average, in the non-degenerate case. It is more practical to consider degeneracy in connection with the symmetry of the system concerned – atoms, molecules or solids – as we will do in the individual cases.

3.5 Quantal vs. classical field description of absorption process

In the preceding section, we described quantum mechanically both the material system and the radiation field. However, the latter can be approximated by a classical electromagnetic field if the field is intense enough that the number of photons n_λ per mode is much greater than unity and the annihilation–creation operators b_λ and b_λ^\dagger can be treated as c-numbers (namely, if their commutator can be neglected compared with n_λ). We will prove here the equivalence of the quantal and classical descriptions for the absorption process in an intense field, partly for later use and partly as another example of the correspondence principle.

Consider the Schrödinger equation for the interacting matter and radiation field:

$$i\hbar d\Psi/dt = (H_M + H_R + H_I)\Psi = (H_0 + H_I)\Psi. \quad (3.5.1)$$

Here $H_I^{(2)}$ of (3.3.8) is omitted while $H_I^{(1)}$ of (3.3.7a), abbreviated as H_I , is rewritten as

$$H_I = \sum_\lambda (b_\lambda h_\lambda + b_\lambda^\dagger h_\lambda^\dagger), \quad (3.5.2)$$

where h_λ and b_λ are operators of the *matter* and *radiation* systems, respectively. Introduce Φ by a unitary transformation:

$$\Psi = S\Phi \equiv \exp(-iH_0 t/\hbar)\Phi. \quad (3.5.3)$$

(Unitarity: $S^\dagger = S^{-1}$ is satisfied since H_0 is an Hermite operator. See the end of Section 2.1 for the definition of the exponential function of an operator H .) It

satisfies the equations

$$i\hbar d\Phi/dt = \exp(i H_0 t/\hbar) H_I \exp(-i H_0 t/\hbar) \Phi \quad (3.5.4)$$

$$= \exp(i H_M t/\hbar) H_I(t) \exp(-i H_M t/\hbar) \Phi, \quad (3.5.4a)$$

where

$$\begin{aligned} H_I(t) &\equiv \exp(i H_R t/\hbar) H_I \exp(-i H_R t/\hbar) = \sum_{\lambda} (b_{\lambda}(t) h_{\lambda} + b_{\lambda}^{\dagger}(t) h_{\lambda}^{\dagger}) \\ &= \sum_{\lambda} \{b_{\lambda} \exp(-i\omega_{\lambda} t) h_{\lambda} + b_{\lambda}^{\dagger} \exp(i\omega_{\lambda} t) h_{\lambda}^{\dagger}\}. \end{aligned} \quad (3.5.5)$$

Let us begin with the *fully quantum-mechanical* description. Consider the transition in the non-perturbed system H_0 with eigenvalues $E_i = E_a + E(\dots, n_{\lambda}, \dots)$ and eigenstates $\Phi_i = |a\rangle | \dots, n_{\lambda}, \dots \rangle$ caused by the interaction H_I with the use of (3.5.4). Putting the expansion:

$$\Phi(t) = \sum_i c_i(t) \Phi_i \quad (3.5.6)$$

into (3.5.4), taking the scalar product of Φ_f with the equation thus obtained, one finds

$$i\hbar dc_f(t)/dt = \sum_j \exp(i\omega_{fj} t) H_{Ifj} c_j(t), \quad \omega_{fj} \equiv (E_f - E_i)/\hbar. \quad (3.5.7)$$

For short times, one can insert the initial condition: $c_j(t) \cong c_j(0) = \delta_{ij}$ on the r.h.s. After integration one obtains

$$c_f(t) = H_{Ifj} \{ \exp(i\omega_{fj} t) - 1 \} / \hbar \omega_{fji}. \quad (3.5.8)$$

For the process in which the matter is excited to the state $|b\rangle$ absorbing one photon from various possible modes λ , one has the total transition probability:

$$\sum_{\lambda} |c_{\lambda}(t)|^2 = \sum_{\lambda} n_{\lambda} |h_{\lambda ba}|^2 [4 \sin^2\{(\omega_{ba} - \omega_{\lambda})t/2\} / \hbar^2 (\omega_{ba} - \omega_{\lambda})^2]. \quad (3.5.9)$$

Considering the photon flux within the solid angle $\Delta\Omega$, the summation on λ can be replaced by the integral:

$$\sum_{\lambda} \rightarrow \int d\omega_{\lambda} \rho(\omega_{\lambda}) \Delta\Omega \quad (3.5.10)$$

where $\rho(\omega)$ is the spectral density of normal modes given by (1.3.7). If t is large enough, the function $[\dots]$ in (3.5.9) has so sharp a peak at $\omega_{\lambda} \sim \omega_{ba}$ that it can be replaced by a δ -function:

$$[\dots] \simeq t(2\pi\hbar^{-2})\delta(\omega_{\lambda} - \omega_{ba})$$

where use has been made of the formula: $\int_{-\infty}^{+\infty} dx \sin^2 x / x^2 = \pi$. The transition rate is then given by $W = \sum_{\lambda} |c_{\lambda}(t)|^2 / t$ which can be readily confirmed to agree with (3.4.6, 3.4.7).

Let us next describe the same process *semi-classically*, and consider the transition $a \rightarrow b$ in the system M with the use of (3.5.4a). In this case, the Hamiltonian, (3.5.5), should be considered as an oscillating (with ω_λ) electromagnetic field driven externally by H_R with *c-number* amplitudes $|b_\lambda| = (n_\lambda)^{1/2}$. The *probability amplitude* of the system M in the state b at time t is then given by

$$c_b(t) = \sum_\lambda h_{\lambda ba} b_\lambda [1 - \exp\{i(\omega_{ba} - \omega_\lambda)t\}]/\hbar(\omega_{ba} - \omega_\lambda). \quad (3.5.11)$$

Since b_λ s with different λ have phase factors which have no correlation in the case of an ordinary light source, one can put

$$\langle b_\lambda b_\lambda^* \rangle_{\text{phase}} = \delta_{\lambda'\lambda} n_\lambda \quad (3.5.12)$$

in calculating $\langle |c_b(t)|^2 \rangle_{\text{phase}}$. The result is the same as that obtained above from the quantal field description, as can be readily confirmed.

3.6 Oscillator strength and the sum rule

For order of magnitude estimation and for comparisons of the rates of different optical transitions, it is convenient to normalize the matrix elements μ_{ba} into dimensionless quantities so as to satisfy the sum rules⁵ to be described below.

Let us start with the commutation relations and the Heisenberg equation of motion for the i th particle of the material system M:

$$\delta_{ij} = (i\hbar)^{-1}[x_i, p_{jx}]_-, \quad p_{jx} = m_j dx_j/dt = m_j i\hbar^{-1}[H, x_j]_-,$$

and take the diagonal matrix element (a, a) of the former with the use of non-diagonal elements of the latter:

$$\begin{aligned} \delta_{ij} &= (i\hbar)^{-1} \sum_b (x_{iab} p_{jxba} - p_{jxab} x_{iba}) \\ &= 2\hbar^{-1} m_j \sum_b \omega_{ba} (x_{iab} x_{jba} + x_{jab} x_{iba}), \end{aligned} \quad (3.6.1)$$

where $|a\rangle$ and $|b\rangle$ are eigenstates of an appropriate Hamiltonian H to be specified later.

If we define the oscillator strength f by

$$f_{iba}^{(x)} \equiv 2\hbar^{-1} m_i \omega_{ba} |x_{iba}|^2 = 2\hbar^{-1} (m_i \omega_{ba})^{-1} |p_{ixba}|^2, \quad (3.6.2)$$

we obtain from (3.6.1) the sum rule for the i th particle:

$$\sum_b f_{iba}^{(x)} = 1. \quad (3.6.3)$$

If i represents one of the n identical particles, the summation of (3.6.1) over i and j gives

$$\sum_b f_{ba}^{(x)} = n \quad (3.6.4)$$

where

$$f_{ba}^{(x)} \equiv 2\hbar^{-1}m\omega_{ba} \left| \left(\sum_i x_i \right)_{ba} \right|^2 = 2\hbar^{-1}(m\omega_{ba})^{-1} \left| \left(\sum_i p_{ix} \right)_{ba} \right|^2. \quad (3.6.5)$$

Equations (3.6.3) and (3.6.4) are the sum rules for one- and n -particle systems, respectively. Only the latter is useful for an *interacting* n -particle system, as will be seen below.

In the case of an atom (or ion) consisting of n electrons and a nucleus with atomic number Z , we first separate their center of mass motion and consider the Hamiltonian H for the electrons relative to the point charge $+Ze$ fixed at the center, and then apply (3.6.4) as an exact relation apart from the extremely small reduced mass correction. Due to the spherical symmetry, one-electron states are characterized by quantum numbers n, l and m , and it is convenient to redefine the oscillator strength by taking the average over $m (= -l, \dots, +l)$ of the $g_l (= 2l + 1)$ -fold degenerate initial states and by summing over m' of the degenerate final states:

$$f_{n'l',nl} \equiv g_l^{-1} \sum_m \sum_{m'} \langle f_{n'l'm',nlm}^{(x)} \rangle, \quad (3.6.6)$$

which no longer depends on the polarization direction (x) of the electromagnetic wave. In the case of the hydrogen atom which is exactly solvable, one can readily confirm the sum rule (3.6.3) (ionization continuum must be included in the final states) for any initial states. Note that there are negative contributions of the downward transitions from initial states 2p and higher.

The outermost electron in alkali atoms, which can be approximately described by an effective one-electron Hamiltonian h , satisfies the sum rule (3.6.3) *approximately*, where a denotes the one-electron state of h . It is only the sum rule (3.6.4) for n electrons with their many-body Hamiltonian H that holds exactly.

In the case of molecules and solids which will be considered in later sections, the interatomic motion remains even after separation of center-of-mass and rotational motions. In the first step of the adiabatic approximation to be discussed in Chapter 4, the electrons are assumed to move around the atoms which are fixed at their equilibrium positions. Then one can apply the sum rule (3.5.4) for the electronic system. In the second step, the atoms are considered to vibrate around their equilibrium positions under the adiabatic potential of the electronic state under consideration, say the ground state. One can then apply a similar sum rule (not shown here explicitly since we have different atomic species) for the atomic motions with their vibrational states. In solids, vibrational and electronic transitions contribute within their respective characteristic ranges of frequency to the dispersion of the dielectric constant. The physical meaning of the oscillator strength and the sum rule will be made more clear in connection with dielectric dispersion described later.

Finally we relate the rate of spontaneous emission, A , to the oscillator strength, f , of the transition. As a typical case, consider an atom in which an electron makes a transition from a p-like excited state to an s-like ground state by emitting a photon of energy $E = \hbar\omega$ into a vacuum spontaneously. Inserting the matrix element of x obtained from (3.6.2) into (3.4.12) where $\mu^2/3$ is to be replaced by $(ex)^2$, one obtains

$$A = (\omega^2 e^2 / 2\pi \epsilon_0 m c_0^3) f = 0.438 \times 10^7 \text{ s}^{-1} E(\text{eV})^2 f, \quad (3.6.7)$$

which will be useful for an order of magnitude estimation.

4

Electronic vs. nuclear motions and the optical spectra of localized electrons in solids

4.1 Born–Oppenheimer approximation

Molecules and condensed matter consist of electrons and nuclei, the latter being ten thousands times *heavier* and hence moving much *more slowly* than the former. According to Born and Oppenheimer¹, it is appropriate to describe such systems in the following two steps: first one solves for the electronic state ℓ with eigenenergy $E_\ell(q)$ for fixed positions q of the nuclei as parameters, and then solves the nuclear motion under the *adiabatic potential* $W_\ell(q)$ which consists of the electronic energy $E_\ell(q)$ and the internuclear potential $U(q)$. Then the eigenenergy $E_{\ell n}$ with quantum number n gives the total energy of the system. In order for this approximation to be valid, however, the electronic state E_ℓ should be a discrete level well separated from other levels. This situation is realized with the electrons of molecules up to medium size or of localized electrons in bulk insulators, and also with the ground electronic state of an insulator which is separated from the excited electronic states by a bandgap.

Let us consider a system consisting of electrons and nuclei with positions \mathbf{r}_j and \mathbf{q}_α , respectively. Since the potential V for the electrons depends on q (representative of all \mathbf{q}_α s) as well as on r (all \mathbf{r}_j s), one can write the total Hamiltonian as

$$H = - \sum_j (\hbar^2/2m) \nabla_j^2 + V(r, q) + U(q) - \sum_\alpha (\hbar^2/2M_\alpha) \nabla_\alpha^2, \quad (4.1.1)$$

where ∇ means the vector differential operator defined below (1.1.4) with respect to the coordinate shown by the suffix j for the j th electron and α for the α th nucleus. (Note (2.1.4) which brings the kinetic energy in (3.3.5) into the first and third terms in (4.1.1).)

One first solves the electronic eigenvalue problem for *fixed* q :

$$\left[- \sum_j (\hbar^2/2m) \nabla_j^2 + V(r, q) - E_\ell(q) \right] \Phi_\ell(r, q) = 0, \quad (4.1.2)$$

and then solves the eigenvalue problem for the nuclear motion with the use of this eigenenergy $E_\ell(q)$ for each electronic state ℓ :

$$\left[- \sum_{\alpha} (\hbar^2/2M_{\alpha}) \nabla_{\alpha}^2 + W_{\ell}(q) - E_{\ell n} \right] \chi_{\ell n}(q) = 0, \quad (4.1.3)$$

$$\text{where} \quad W_{\ell}(q) \equiv E_{\ell}(q) + U(q). \quad (4.1.4)$$

The wave function of the total system is then given by

$$\Psi_{\ell n}(r, q) \cong \Phi_{\ell}(r, q) \chi_{\ell n}(q). \quad (4.1.5)$$

To see the accuracy of this approximation, we take the scalar product of $\chi_{\ell n}$ with eq. (4.1.2) and the scalar product of Φ_{ℓ} with eq. (4.1.3) and add them. Then the discrepancy of $E_{\ell n}$ and $\Psi_{\ell n}$ from the exact eigenenergy and eigenfunction of H is given by

$$\begin{aligned} (H - E_{\ell n})\Psi_{\ell n} &= - \sum_{\alpha} (\hbar^2/2M_{\alpha}) [\nabla_{\alpha}^2, \Phi_{\ell n}]_- \chi_{\ell n}(q) \\ &= - \sum_{\alpha} (\hbar^2/2M_{\alpha}) [2(\nabla_{\alpha} \Phi_{\ell n}) \cdot (\nabla_{\alpha} \chi_{\ell n}) + (\nabla_{\alpha}^2 \Phi_{\ell n}) \chi_{\ell n}]. \end{aligned} \quad (4.1.6)$$

Taking derivatives with respect to q of $\Phi(r, q)$ and $\chi(q)$ amounts, in order of magnitude, to dividing by the electronic and nuclear de Broglie wavelengths λ_{el} and λ_{nuc} (to be defined later in this section), respectively. (Note that the *electronic* wave function Φ depends on q only through the relative coordinates $(\mathbf{r}_j - \mathbf{q}_{\alpha})$.) Therefore, the first and second terms of (4.1.6) are of order of magnitude $(\lambda_{\text{nuc}}/\lambda_{\text{el}})$ and $(\lambda_{\text{nuc}}/\lambda_{\text{el}})^2$ times the nuclear kinetic energy given by the first term in $[\cdots]$ of (4.1.3).

Let us expand the adiabatic potential around its minimum point which naturally depends on ℓ : $W_{\ell}(q) = W_{\ell}(q_{0\ell}) + M\omega_{\ell}^2(q - q_{0\ell})^2/2 + \cdots$. Since the ω_{ℓ} s for different ℓ s are of the same order of magnitude, they will be represented by a common ω for the estimation below. In the harmonic approximation, the vibrational energy quantum is given by

$$\hbar\omega \sim \hbar^2/2M\lambda_{\text{nuc}}^2 \sim M\omega^2\lambda_{\text{nuc}}^2/2, \quad (4.1.7)$$

whereas the electronic energy E_{el} defined as the difference of two adjacent energy levels is related to λ_{el} by

$$E_{\text{el}} \sim \hbar^2/2m\lambda_{\text{el}}^2. \quad (4.1.8)$$

Note that $W_{\ell}(q)$ contains the term $E_{\ell}(q)$ whose q dependence comes from that of $V(|\mathbf{r}_j - \mathbf{q}_{\alpha}|)$ in (4.1.2) as mentioned before. Since the internuclear bonds come from the energy of the electrons participating in the bonds, a nuclear displacement as large as the electron de Broglie wavelength would change the electronic state completely, as is seen from the typical situation depicted in Fig. 4.1. Thus the

of (3.4.13, 3.4.14), by the statistical average of energy of *vertical transition* (the Franck–Condon principle, see also Fig. 4.1):

$$B(E) = (\pi/\epsilon\hbar^2) \int dq w_g(q) |\mu_{\lambda eg}(q)|^2 \delta[W_e(q) - W_g(q) - E], \quad (4.2.1)$$

where

$$\mu_{\lambda eg}(q) \equiv \int dr \Phi_e^*(r, q) \mathbf{e}_\lambda \cdot \boldsymbol{\mu}(r) \Phi_g(r, q) \quad (4.2.2)$$

is the component, in the direction \mathbf{e}_λ of polarization of the incident light, of the q -dependent electronic transition dipole moment and

$$w_g(q) \equiv Z_g^{-1} \exp[-\beta W_g(q)], \quad Z_g \equiv \int dq \exp[-\beta W_g(q)] \quad (4.2.3)$$

is the Boltzmann distribution of the vibrational coordinate q . In the above, q is to be understood to represent multi-dimensional coordinates, $\{q_i\}$.

4.3 Perturbation theory in electron–vibration interaction

Let us expand $V(r, q)$, the only term which depends both on r and q in (4.1.1), around q_0 (to be specified later) in a power series

$$V(r, q) = V(r) + \sum_i V_i(r)(q_i - q_{i0}) + \cdots, \quad (4.3.1)$$

and also expand the eigenenergies $E_\ell(q)$ and the wave functions $\Phi_\ell(r, q)$ in the same way. Defining the electron–vibration interaction H_I by the q -dependent term (the second term on the r.h.s.) of (4.3.1), we apply the perturbation theory to solve the eigenequation (4.1.2).² The coefficients of the zeroth and the first powers of $(q - q_0)$ are then given by

$$[H_0 - E_\ell]\Phi_\ell = 0, \quad H_0 \equiv -\sum_j (\hbar^2/2m)\nabla_j^2 + V(r), \quad (4.3.2)$$

$$[V_i - E_{\ell i}]\Phi_\ell + [H_0 - E_\ell]\Phi_{\ell i} = 0. \quad (4.3.3)$$

The scalar product of $\Phi_{\ell'}$ with (4.3.3) gives

$$V_{i\ell'\ell} - \delta_{\ell'\ell} E_{\ell i} + [E_{\ell'} - E_\ell](\Phi_{\ell'}, \Phi_{\ell i}) = 0.$$

By putting $\ell' = \ell$, one obtains the coefficient of the first-order energy:

$$E_{\ell i} = V_{i\ell\ell}. \quad (4.3.4)$$

From the non-diagonal terms $\ell' \neq \ell$, one obtains the expansion coefficient of the first-order wave function:

$$c_{i\ell\ell'} \equiv (\Phi_{\ell'}, \Phi_{\ell i}) = V_{i\ell'\ell}/[E_\ell - E_{\ell'}]. \quad (4.3.5)$$

As for the coefficient $c_{i\ell\ell}$, one obtains, from the first-order term in the expansion of the normalization condition: $\delta_{\ell\ell'} = (\Phi_\ell, \Phi_{\ell'})$, the relation: $c_{i\ell\ell} + c_{i\ell\ell}^* = 0$, namely, its real part vanishes while the imaginary part is arbitrary and hence can be taken to be zero. Hence $c_{i\ell\ell} = 0$. It has been assumed in the above that the state ℓ is not degenerate.

In studying the absorption spectrum of the transition $g \rightarrow e$, it is convenient to take q_{i0} as the minimum point of $W_g(q)$ and to choose their origins as $q_{i0} = 0$ and $W_g(0) = 0$. One can then expand

$$W_g(q) = \sum_i \omega_i^2 q_i^2 / 2 + \cdots, \quad (4.3.6)$$

$$\begin{aligned} W_e(q) &= E_a - \sum_i \omega_i^2 \Delta_i q_i + W_g(q) + \cdots \\ &= E_0 + \sum_i \omega_i^2 (q_i - \Delta_i)^2 / 2 + \cdots, \end{aligned} \quad (4.3.7)$$

$$E_R \equiv E_a - E_0 = \sum_i \omega_i^2 \Delta_i^2 / 2, \quad (4.3.8)$$

where the q_i s are normal modes so chosen as to diagonalize the quadratic term of the potential energy $W_g(q)$ as shown in (4.3.6) and the kinetic energy as $\sum_i p_i^2 / 2$ with unit mass. The linear interaction term in (4.3.7) are given by

$$V(q) = - \sum_i \omega_i^2 \Delta_i q_i \equiv \sum_i (V_{iee} - V_{igg}) q_i \quad (4.3.9)$$

with the use of (4.3.4) for the relative displacement of the equilibrium position.

In the harmonic approximation where the third- and higher-order terms \dots in (4.3.6) are neglected, the terms \dots in (4.3.7) consist of the second- and higher-order perturbations in H_I (although the second-order terms are already included in the ground-state harmonic potential $W_g(q)$). We shall hereafter refer to this approximation (the neglect of terms \dots in (4.3.7) and (4.3.6)) as *linear interaction* of electrons with the *harmonic* vibration (see Fig. 4.1). With further use of the Condon approximation which neglects the q dependence of the electronic transition dipole moment, one readily finds from (4.2.1) that the normalized absorption spectrum has a Gaussian lineshape:

$$F(E) = (2\pi D^2)^{-1/2} \exp[-(E - E_a)^2 / 2D^2], \quad (4.3.10)$$

$$\langle\langle E \rangle\rangle \equiv \int E F(E) dE = E_a \quad (4.3.11)$$

$$\langle\langle E^2 \rangle\rangle \equiv \int E^2 F(E) dE \equiv D^2 = \langle\langle (V(q))^2 \rangle\rangle = 2E_R k_B T. \quad (4.3.12)$$

4.4 Relaxation of nuclear positions after electronic transition

We are now in a position to consider what takes place in the nuclear positions after optical excitation of the electronic state. According to the Franck–Condon principle, the nuclei immediately after electronic excitation have the same positions $\langle\langle q_i \rangle\rangle = 0$ and momenta $\langle\langle p_i \rangle\rangle = 0$ as they had before the excitation. We can resort to classical mechanics in describing the subsequent motions of nuclei so long as we take the harmonic approximation for the adiabatic potential, because quantum mechanics gives the same result (on the expectation values of the position coordinates) as classical mechanics according to (2.2.27 to 2.2.28). Starting from the Franck–Condon state mentioned above, the nuclei move under the potential $W_e(q)$ given by (4.3.7). Hence, each normal mode makes a harmonic oscillation around its new equilibrium position Δ_i with amplitude Δ_i :

$$q_i(t) = \Delta_i(1 - \cos \omega_i t). \quad (4.4.1)$$

The quantity that is spectroscopically observable is the vertical transition energy $h\nu = W_e(q) - W_g(q) = E_a + V(q)$ (see (4.3.7, 4.3.8)) between the adiabatic potentials according to the Franck–Condon principle, rather than the individual motion of each normal mode. It varies rapidly, sub-picosecond, the typical period of the vibrational motions. However, the time-resolved observation of weak light emitted during these motions is well within the scope of recent spectroscopic techniques which operate in the femtosecond regime. Since $V(q)$ is a linear function of the q_i s as given by (4.3.8), it is most straightforward to take it as one of the new coordinate axes.

For this purpose we first make a canonical transformation: $q_i \rightarrow q_i' \equiv \omega_i q_i$ and $p_i \rightarrow p_i' \equiv \omega_i^{-1} p_i$ by which the potential W_g is brought into an invariant form at the cost of bringing back the kinetic energy K to a non-invariant form:

$$W_g = \sum_i q_i'^2/2, \quad K = \sum_i \omega_i^2 p_i'^2/2. \quad (4.4.2)$$

Second, we make an orthogonal transformation $\{q_i'\} \rightarrow \{Q_\alpha\}$ (α is not to be confused with α in Section 4.1) such that Q_1 , the first of the new coordinates, is proportional to $V(q)$ given by (4.3.8):

$$V(q) = - \sum_i \omega_i \Delta_i q_i' = -c Q_1, \quad (4.4.3)$$

$$c^2 = \sum_i \omega_i^2 \Delta_i^2 = 2E_R. \quad (4.4.4)$$

Q_1 is called the *interaction mode* since it is proportional to the interaction $V(q)$ except for a constant c to be defined as a coupling coefficient, while the other Q_α s do not appear in the interaction. The interaction mode was first introduced to describe the Jahn–Teller degenerate excited states – the optical transition thereto and

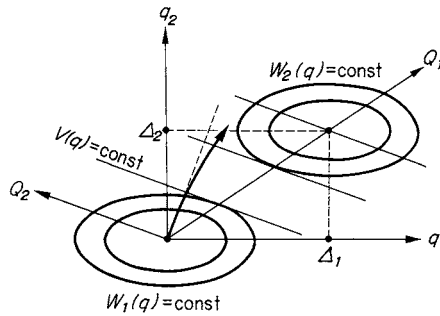


Fig. 4.2 Contours of adiabatic potentials for the ground and excited electronic states in the configuration coordinate space (q_1, q_2, \dots) , and the interaction mode Q_1 upon which the relaxational motion (thick arrow) is to be projected so long as one is concerned only with the optical spectra.³

the lattice relaxation therefrom – in a tractable way with the use of the configuration coordinate model³ as will be described in detail in Section 4.7. In that case, several interaction modes with a few coupling coefficients are needed depending on the symmetry of the system. The concept is still useful in the present case of a non-degenerate excited state with a single interaction mode with its one coupling coefficient. It is to be noted that the combined transformation $\{q_i\} \rightarrow \{Q_\alpha\}$ is a non-orthogonal one as shown schematically in Fig. 4.2. Then one can rewrite the adiabatic potentials in terms of the new coordinates:

$$W_g(Q) = \sum_{\alpha} Q_{\alpha}^2/2, \quad (4.4.5)$$

$$W_e(Q) = E_a - cQ_1 + W_g(Q) = E_0 + (Q_1 - c)^2/2 + \sum_{\alpha \geq 2} Q_{\alpha}^2/2. \quad (4.4.6)$$

The quadratic form K is no longer diagonal in the new momenta P_{α} , making the situation complicated. However, since we are concerned only with the adiabatic potential along the Q_1 -axis, the axis relevant to spectroscopy, we can pursue the temporal behavior explicitly given by

$$cQ_1(t) = \sum_i \omega_i^2 \Delta_i^2 (1 - \cos \omega_i t) = 2 \int d\omega s(\omega) \hbar \omega (1 - \cos \omega t), \quad (4.4.7)$$

$$s(\omega) \equiv \sum_i S_i \delta(\omega_i - \omega), \quad S_i \equiv (\omega_i \Delta_i^2 / 2\hbar). \quad (4.4.8)$$

The integral expression in (4.4.7), with the use of *coupling function* $s(\omega)$ as defined by (4.4.8),⁴ is convenient in studying a localized electron at a defect in a solid and, in particular, the different roles of localized and delocalized modes of the lattice vibrations which constitute the discrete and continuous spectra of $s(\omega)$.

In the absence of localized modes, $s(\omega)$ is a *continuous* function of ω with an upper bound ω_0 common to the host crystal. All modes are extended, with Δ_i being

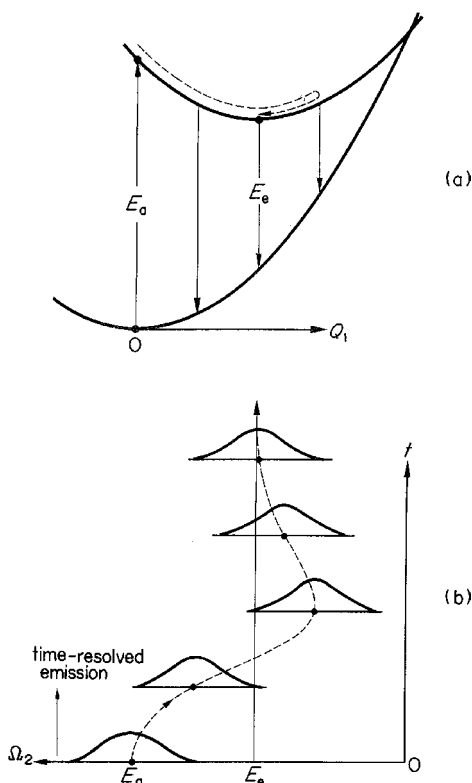


Fig. 4.3 (a) One-dimensional configuration coordinate model with interaction mode Q_1 upon which the relaxational motion is projected.^{3,4} (b) Expected time-resolved emission spectra. (See also Section 11.5 and compare with Fig. 11.6.)

on the order of $N^{-1/2}$ (N : the total number of atoms in the crystal) since they interact with the localized electron only at several atoms in its neighborhood. The integral (4.4.7) tends to $cQ_1(\infty) = c^2$ as a result of dephasing among oscillations with different ω , though with some to-and-fro motion on its way. The emission spectrum after photoexcitation is expected to change during the relaxation ($\sim 10^{-13}$ s, as shown schematically in Fig. 4.3,⁵ although the intensity of this *hot luminescence* is extremely small compared with the ordinary luminescence from the *relaxed excited state* at $Q_1 = c$ where the system stays much longer (during the natural lifetime which is of the order of 10^{-8} s, see (3.6.7)). All other Q_α s ($\alpha \neq 1$) make oscillative (but not periodic) motions with amplitude of the order of $N^{-1/2}$ as t tends to ∞ . These motions constitute the heat energy dispersed over the entire crystal, the energy originating from the local relaxation energy: $E_R = c^2/2$. When one resorts to the one-dimensional *configuration coordinate* (C. C.) model, as shown schematically in Figs. 4.1 and 4.3, which has for many years been very useful in describing the optical transitions (inclusive of emission) as well as the subsequent

relaxations, the abscissa should be considered to represent the interaction mode described above. The ordinate of the relaxed state should be interpreted as the *free energy* ($F = U - TS$ in conventional thermodynamics) of the electron localized in the lattice since the energy, E_R , lost during relaxation is converted to the heat energy (TS , where S is the entropy) of the lattice.

However, the localized normal modes, if they exist, contribute *discontinuous* (line) spectra to $s(\omega)$, and, according to (4.4.7), $Q_1(t)$ continues to oscillate until the anharmonic terms in $W_e(q)$, which have so far been neglected, dissipate their energy as heat.

4.5 Vibrational structures in absorption spectra

Let us now treat the nuclear motions quantum mechanically. If we confine ourselves to the Condon approximation, the transition dipole moment defined in (3.4.11) is factorized into the electronic and vibrational ones:

$$\mu_{en, gm} = (\Phi_e, \mu \Phi_g)(\chi_{en}, \chi_{gm}). \quad (4.5.1)$$

Apart from a constant factor, the absorption spectrum at temperature T is given by

$$F(E) = Z \sum_m \exp(-\beta E_{gm}) \sum_{n'} \delta(E_{en} - E_{gm} - E) |(\chi_{en}, \chi_{gm})|^2, \quad (4.5.2)$$

$$\text{with } Z \equiv \sum_m \exp(-\beta E_{gm}) \quad (4.5.3)$$

where $F(E)$ is normalized as is obvious from the completeness of the eigenfunctions $\{\chi_{en}\}$.

According to the model of linear interaction of the electrons with the multi-dimensional harmonic vibrations as introduced by (4.3.6) to (4.3.8), the eigenstates are specified by $(gm) \equiv (g; \dots, m_i, \dots)$ and $(en) \equiv (e; \dots, n_i, \dots)$, and the overlap of their vibrational states are given by the product over all normal modes i :

$$(\chi_{en}, \chi_{gm}) = \prod_i \int dq_i \chi_{ni}(q_i - \Delta_i) \chi_{mi}(q_i). \quad (4.5.4)$$

Let us confine ourselves to the one-dimensional harmonic oscillator for the moment. At $T = 0$, only the lowest vibrational state: $m = 0$ is populated. With the use of the wave functions (2.2.25, 2.2.26), partial integration and recursion formula, one can rewrite a typical integral of (4.5.4) for $m = 0$ as

$$\begin{aligned} c_n(\Delta) &\equiv \int dq \chi_n(q - \Delta) \chi_0(q) \\ &= (2n)^{-1/2} \int dq' \chi_{n-1}(q') \{q'/a + a(d/dq')\} \chi_0(q' + \Delta) \\ &= (2n)^{-1/2} (-\Delta/a) c_{n-1}(\Delta) = (n!)^{-1/2} (-\Delta/\sqrt{2}a)^n c_0(\Delta) \end{aligned} \quad (4.5.5)$$

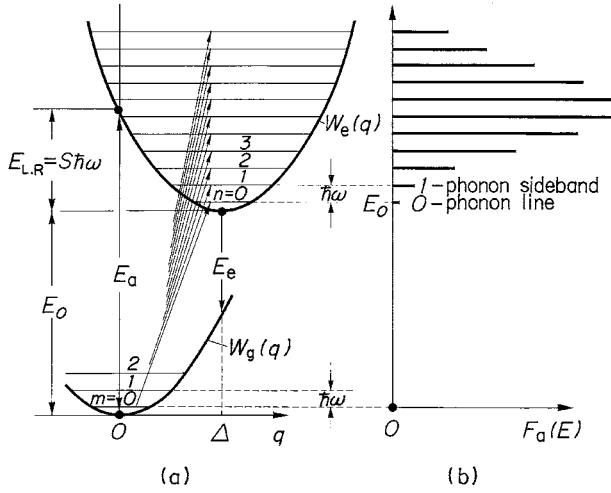


Fig. 4.4 Poisson distribution of phonon lines (b) in the electronic absorption spectra for the model (a) of linear electron-phonon interaction for a one-dimensional harmonic lattice vibration (see eq. (4.5.6)).

where $a \equiv (\hbar/\omega)^{1/2}$ (see eq. (2.2.3) for unit mass). Hence, the normalized intensities of the vibrational lines (of the transition $n \leftarrow 0$) are given by the Poisson distribution:

$$I_n = |c_n(\Delta)|^2 = \exp(-S) \times S^n / n! \quad \left(\sum_n I_n = 1 \right), \quad (4.5.6)$$

$$\text{where } S \equiv \omega \Delta^2 / 2\hbar = E_R / \hbar\omega \quad (4.5.7)$$

(see Fig. 4.4), and the normalized absorption spectrum by

$$F(E) = \exp(-S) \times \sum_n (S^n / n!) \delta(E_0 + n\hbar\omega - E). \quad (4.5.8)$$

By putting the intensity ratio: $I_n / I_{n-1} = (S/n) \cong 1$, one finds that the maximum of the Poisson distribution (4.5.6) is at $\sim n - (1/2) \cong S - (1/2)$. On the other hand, the mean value $\langle n \rangle$ and its fluctuation Δn defined by $(\Delta n)^2 \equiv \langle (n - \langle n \rangle)^2 \rangle = \langle n^2 \rangle - \langle n \rangle^2$ of the distribution (4.5.6) are found to be equal to S and \sqrt{S} , respectively. The former result gives the mean energy of the optical transition: $\langle E \rangle = E_0 + \langle n \rangle \hbar\omega = E_0 + S\hbar\omega = E_0 + E_R = E_a$ which is nearly at the above-mentioned peak $E_p \cong E_0 + (S - 1/2)\hbar\omega$, the small discrepancy being due to the positive skewness of the Poisson distribution characterized by a longer tail at higher n . Here S represents the *coupling strength* of the phonons with the localized excitation.

The generalization to the multi-dimensional case is as follows.^{2,6} Restoring the indices of the normal modes, i , one can write the normalized absorption spectrum

at $T = 0$ as

$$F(E) = \exp\left(-\sum_i S_i\right) \sum_{n_1} \sum_{n_2} \cdots \prod_i (S_i^{n_i}/n_i!) \delta\left(E_0 + \sum_i n_i \hbar\omega_i - E\right). \quad (4.5.9)$$

Here the δ -function with its coefficients indicates that energy quanta $\hbar\omega_i$ are emitted simultaneously with the electronic excitation energy E_0 , with the probability S_i . A vibrational quantum of interatomic motion is often called a “phonon”. This terminology originates in solids where the acoustic mode of lattice vibrations with long wavelengths correspond to sound waves, but is now used in more general systems.

The complicated summation in (4.5.9) is greatly simplified by taking its Fourier transform, the so-called *generating function*.^{2,7}

$$f(t) \equiv \int_{-\infty}^{+\infty} dE F(E) \exp(-iEt/\hbar) = \exp[-iE_0t/\hbar - S + S(t)], \quad (4.5.10)$$

where

$$S(t) \equiv \sum_i S_i \exp(-i\omega_i t) = \int_0^\infty d\omega s(\omega) \exp(-i\omega t), \quad (4.5.11)$$

$$S \equiv S(0) = \sum_i S_i = \int_0^\infty d\omega s(\omega). \quad (4.5.12)$$

These expressions provide us with an insight to their extension to finite temperature. One has to multiply each term with suffix i in $S(t)$, representing the *spontaneous emission* of a phonon as mentioned above, by $n_\beta(\omega_i) + 1$ so as to include the *induced emission* which is proportional to $n_\beta(\omega_i) = [\exp(\beta\hbar\omega_i) - 1]^{-1}$, the average number of phonons in thermal equilibrium. (See the arguments in Chapter 3 where the quantum concerned was a photon.) At the same time, one has to consider the phonon absorption term, with the $-$ (minus) sign replaced by a $+$ (plus) sign in the exponent of (4.5.11), and multiplied by the factor $n_\beta(\omega_i)$, instead of $n_\beta(\omega_i) + 1$ (see Refs. [2, 6 or 7] for the proof). With the use of (4.4.8), one obtains:

$$f_T(t) = \exp[-iE_0t/\hbar - S_T + S_+(t) + S_-(t)], \quad (4.5.10a)$$

where

$$S_\pm(t) \equiv \int_0^\infty d\omega \{n_\beta(\omega) + 1, \text{ or } n_\beta(\omega)\} s(\omega) \exp(\mp i\omega t), \quad (4.5.11a)$$

$$S_T \equiv S_+(0) + S_-(0) = \int_0^\infty d\omega \coth(\hbar\omega/2k_B T) s(\omega). \quad (4.5.12a)$$

The absorption spectrum is obtained as the inverse Fourier transform of the generating function:

$$F(E) = (2\pi\hbar)^{-1} \int_{-\infty}^{\infty} dt f(t) \exp(itE/\hbar). \quad (4.5.13)$$

Note that the zeroth and first-order terms in the power series expansion in $S_+(t)$ of (4.5.10a) at $T = 0$ gives the following terms for the absorption spectrum:

$$F_0(E) = \exp(-S) \delta(E - E_0), \quad (4.5.14)$$

$$F_1(E) = \exp(-S) s((E - E_0)/\hbar)/\hbar. \quad (4.5.15)$$

Equation (4.5.14) represents the zero-phonon line, and (4.5.15) the one-phonon sideband on its higher-energy side reflecting the spectral coupling $s(\omega)$ of the interacting phonons. Likewise, the two-phonon sideband is obtained from the second-order terms in the power series expansion in the form of a convolution:

$$F_2(E) = \exp(-S)(2!)^{-1} \int dE' s((E - E' - E_0)/\hbar) s(E'/\hbar)/\hbar. \quad (4.5.16)$$

At finite temperature the sidebands of the phonon absorption appear on the low-energy side of the zero-phonon line, and the mixed bands of phonon emission and absorption on both sides of it, all of which are given in convolutions of $s(\omega)$ (see Fig. 4.5).

For high temperature and/or strong electron-phonon coupling where the mean square root width D of the absorption spectrum, given by

$$D^2 \equiv \langle \langle V(q)^2 \rangle \rangle_T = \int d\omega \cot(\hbar\omega/2k_B T) (\hbar\omega)^2 s(\omega), \quad (4.5.17)$$

is large compared to the typical phonon energy $\hbar\omega$, only the region with small t of $f_T(t)$ contributes to $F(E)$. Expanding the exponent of (4.5.10a) in powers of t , one obtains $-iE_a t/\hbar - (Dt)^2/2\hbar^2$, resulting in the Gaussian lineshape (4.3.10) with dispersion $\langle \langle (E - E_a)^2 \rangle \rangle = D^2$. In such a situation where $S_T \sim (D/\hbar\omega)^2$ is much larger than unity, the intensity of the zero-phonon line, $\exp(-S_T)$, is negligibly small, and successive phonon sidebands, if continuous, can hardly be resolved from each other. Note that D tends to eq. (4.3.12) in the high-temperature limit, as it should.

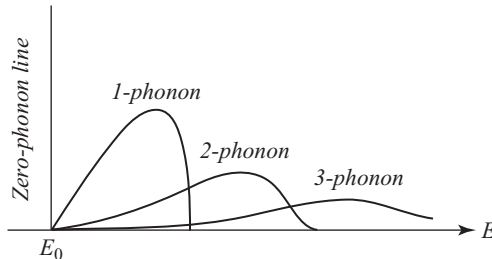


Fig. 4.5 The zero-phonon line followed by one-, two-, and three-phonon sidebands (schematic) of the electronic absorption spectra (see eqs. (4.5.14 to 4.5.16)).

4.6 Emission vs. absorption spectra

While relaxation after photo-excitation takes place within the order of a picosecond, spontaneous emission from the excited state takes much longer than nanoseconds according to (3.6.7). It is therefore quite reasonable to suppose that thermal equilibrium among vibrational states within the excited electronic state is established well before spontaneous emission takes place. Hence, the emission spectra $A(E)$ can be derived in the same way as the absorption spectra. For instance, one obtains the classical approximation:

$$A(E) = (E^3/\pi\epsilon\hbar^4c^3) \int dq w_e(q) |\mu_{\lambda_{eg}}(q)|^2 \delta[W_e(q) - W_g(q) - E] \quad (4.6.1)$$

where $w_e(q)$ denotes the Boltzmann distribution within the excited electronic state as obtained by replacing g by e in (4.2.3). One can readily confirm that (4.2.1) and (4.6.1) satisfy the relation (3.2.3) as they should.

With the model of linear interaction with the harmonic vibrations presented in Section 4.3, the emission spectrum in the classical approximation has a Gaussian lineshape (apart from the factor E^3) with the same dispersion D^2 as in (4.3.10) but with its peak E_e shifted by the Stokes shift:

$$E_e = E_0 - E_R = E_a - 2E_R \quad (4.6.2)$$

as shown schematically in Figs. 4.1 and 4.3. After emission, the nuclei return to the original positions they were at before excitation, again with the relaxation energy E_R . Quantum-mechanical treatment of the same model gives the generating function which is obtained by replacing $\exp(\mp i\omega t)$ by $\exp(\pm i\omega t)$ in (4.5.11a), corresponding to the fact that the simultaneous emission of phonons contributes now to the low energy shift of the emitted photon. Its inverse Fourier transformation, which gives the emission spectrum apart from the factor E^3 , turns out to be a mirror image of the absorption spectrum with the zero-phonon line as the point of reflection.

It is to be noted that the reflection symmetry is valid only for the model of linear interaction with harmonic vibrations, whereas relation (3.2.3) is valid for any model provided that thermal equilibrium is established within the excited-state subspace before the emission. The time-dependent weak emission during the relaxation, as depicted in Fig. 4.3, makes by itself a negligible contribution to this relation because it is only one channel of relaxation (starting from a particular initial state – the Franck–Condon (F.C.) state) towards thermal equilibrium among innumerable channels. (In this connection one is to be reminded of the ergodic theory relating the reversible mechanical law to the irreversible processes which finally establish thermodynamical equilibrium.)

4.7 Interaction modes in degenerate excited states

As we have seen in Sections 4.3 and 4.4, the absorption spectrum of a molecule or a localized electron in solids with non-degenerate ground and excited states can be described, in the adiabatic limit, with the use of the one-dimensional configuration coordinate (C.C.) model, under the simplified assumption of *linear interaction* of the electron with the *harmonic vibrations* of the atoms. This assumption allowed us to define the *interaction modes* as that linear combination of the normal modes which is proportional to the interaction Hamiltonian itself. Although the interaction mode is not a normal mode, the lineshape of the absorption spectrum in the classical limit can be described by the C.C. model with this interaction mode alone, the other modes playing the role of a heat reservoir which accepts the energy released during relaxation of the interaction mode toward its new equilibrium position (see Fig. 4.3).

However, molecules and lattice defects in solids⁸ with which we are familiar very often have high symmetry, so that the optically allowed excited states are degenerate even if the totally symmetric ground state is not. Consider a lattice defect with cubic symmetry (O_h). A well-known example is the F center in alkali halides⁹ which consists of a halide-ion vacancy with a captured electron. It has a strong absorption band (called the F band) in the visible region, resulting in complementary coloring of otherwise transparent crystals – the origin of the German name: *Farbzentren*. The ground state of the captured electron is a totally symmetric A_{1g} state (similar to the atomic s state) which is non-degenerate, while the excited states responsible for the strong absorption band are triply-degenerate T_{1u} states (similar to the atomic p state). Another example known for its phosphorescence is the impurity center with a heavy-metal ion with an outermost shell of $(ns)^2$ configuration (such as In^+ , $Sn^{2+}(n = 5)$, Tl^+ , $Pb^{2+}(n = 6)$) substituting for an alkali ion. The optical absorption spectra of this center (sometimes called a Tl^+ -like center, representatively) consist, in order of increasing energy, of a spin-orbit-allowed A band, a vibrationally-allowed B band (both originating from the spin triplet excited state) and a dipole-allowed C band. We shall be mainly concerned here with the C band which can be approximately considered as a one-electron transition: $ns \rightarrow np$, similar to the F band. However, the electron responsible for the optical transition is *well localized within the impurity atom* in the Tl^+ -like center, but *spread to neighboring host ions* in the F center.

According to the well-known Jahn-Teller theorem,¹⁰ the orbitally degenerate states are stabilized in the distorted configurations in which the degeneracy is removed. This is caused by the interaction of the electron with modes (patterns) of atomic displacements with lower symmetries. For instance, T_{1u} states which transform like a vector (x, y, z) are subject to a linear interaction with totally symmetric A_{1g} modes represented symbolically by Q_1 , two-dimensional E_g modes

(Q_2, Q_3) transforming like $((\sqrt{3}/2)(x^2 - y^2), -(1/2)(x^2 + y^2) + z^2)$, and three-dimensional T_{2g} modes (Q_4, Q_5, Q_6) transforming like $(yz, zx, xy)^{11}$ through the Hamiltonian H_I of the matrix form:

$$H_I = \begin{pmatrix} \begin{matrix} x & y & z \\ \cdots & \cdots & \cdots \end{matrix} \\ \begin{matrix} a'Q_1 + b'[(\sqrt{3}/2)Q_2 - (1/2)Q_3] & c'(\sqrt{3}/2)Q_6 & c'(\sqrt{3}/2)Q_5 \\ c'(\sqrt{3}/2)Q_6 & a'Q_1 + b'[-(\sqrt{3}/2)Q_2 - (1/2)Q_3] & c'(\sqrt{3}/2)Q_4 \\ c'(\sqrt{3}/2)Q_5 & c'(\sqrt{3}/2)Q_4 & a'Q_1 + b'Q_3 \end{matrix} \end{pmatrix} \quad (4.7.1)$$

Here Q_1, \dots, Q_6 are linear combinations of all normal modes $\{q_{1i}\}, \dots, \{q_{6j}\}$ of the corresponding symmetries with coefficients $\{\gamma_{1i}\}, \dots, \{\gamma_{6j}\}$ common within each of *irreducible representations* A_{1g}, E_g and T_{2g} . By appropriate transformation of the C.C.s within each irreducible representation, one can now reduce the N -dimensional C.C. space to a six-dimensional one with three independent coupling coefficients a', b', c' , instead of a one-dimensional one with one coupling coefficient c as was done in Section 4.4 for the non-degenerate case.

Adding three eigenvalues of the energy matrix (4.7.1) to the relevant part of the lattice potential energy, $(Q_1^2 + \dots + Q_6^2)/2$, one obtains three sheets (surfaces) of adiabatic potentials in the six-dimensional C.C. space. For any direction in this space, the eigenvalues of H_I are linear functions of the radial coordinate Q and hence each branch of the adiabatic potentials has a minimum as in the non-degenerate case of Section 4.4. Considering all directions, one can find a few (how many is not obvious) equivalent minima of the adiabatic potentials in the six-dimensional C.C. space. This is the Jahn-Teller effect.

Does this splitting of the adiabatic potential in the excited state cause a *split lineshape* of the absorption band? To answer this question, let us first consider the cases in which only one among a, b, c is nonvanishing. The first case with $a \neq 0, b = c = 0$ is trivial since all eigenvalues are equal to aQ_1 (degeneracy is *not* removed), giving three identical adiabatic potential surfaces with only one minimum $-E_R = -a'^2/2$ at $(Q_1, Q_2, \dots) = (-a', 0, \dots)$, which is a totally symmetric configuration. One obtains a Gaussian lineshape with dispersion $D = a'(k_B T)^{1/2}$. In the second case with $b' \neq 0, a = c = 0$ the degeneracy *is* certainly removed. Taking the third diagonal element as an eigenvalue, one finds the adiabatic potential $W_e(Q_2, Q_3) = E_a + b'Q_3 + (Q_2^2 + Q_3^2)/2$ has a minimum $-E_R = -b'^2/2$ at $(Q_2, Q_3) = (0, -b')$. Since Q_2 does not interact, one obtains a Gaussian lineshape with dispersion $D = b'(k_B T)^{1/2}$. Two other sheets of the adiabatic potential are obtained simply by rotating this sheet by 120° and 240° on the (Q_2, Q_3) plane. Hence, the lineshape is a simple Gaussian in spite of the removal of degeneracy for given (Q_2, Q_3) . The situation is different in the case of $a = b = 0, c \neq 0$ where one obtains a *split lineshape* reflecting in fact the removal of degeneracy,³ as will be shown below.

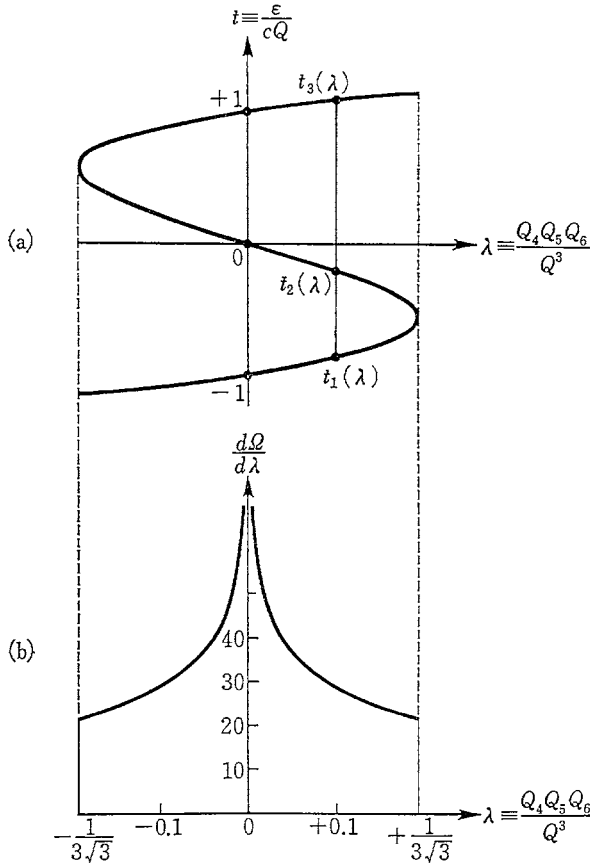


Fig. 4.6 (a) Three solutions of eq. (4.7.2), and (b) differential solid angle ($d\Omega/d\lambda$) given by eq. (4.7.6), as functions of λ .³

Introducing the radius Q with directional cosines $\ell = Q_4/Q$, $m = Q_5/Q$ and $n = Q_6/Q$ in the three-dimensional space (Q_4, Q_5, Q_6) , one can write the eigenvalues of H_1 as $c'Q t_i$ where $t_i(\lambda)$ s are roots of the cubic equation

$$t^3 - (3/4)t - (3^{3/2}/4)\lambda = 0, \quad \text{where } \lambda \equiv \ell mn. \quad (4.7.2)$$

Three roots t_1, t_2 and t_3 in increasing order are illustrated in Fig. 4.6(a), as functions of λ which varies between $-\gamma$ and $+\gamma$ where $\gamma = 3^{-3/2}$. Since $(t_1)_{\min} = -1$ appears at $\lambda = -\gamma$, the adiabatic potential has minima, $-E_R = -c'^2/2$ at $Q = c'$ in four equivalent directions with one or two of (ℓ, m, n) being $+3^{-1/2}$ and the others $-3^{-1/2}$. We have defined the interaction coefficients (a', b', c') in (4.7.1) in such a way that the corresponding lattice relaxation energy E_R is given by $a'^2/2, b'^2/2, c'^2/2$, respectively, in accordance with the non-degenerate case (4.4.4). They are related with (a, b, c) used in Refs. [3, 13, 20–22] by $a = \sqrt{2}a', b = \sqrt{3}b'/2, c = \sqrt{3}c'/2$.

The normalized lineshape for the $A_{1g} \rightarrow T_{1u}$ transition in the Condon approximation is then given by

$$F(E) = (1/3) \sum_i \int d\Omega \int_0^\infty Q^2 dQ (2\pi k_B T)^{-3/2} \times \exp[-Q^2/2k_B T] \delta[E - \{E_a + c'Q t_i(\lambda)\}], \quad (4.7.3)$$

where $d\Omega = (d\Omega/d\lambda) d\lambda$ is the differential solid angle of the directional region in which ℓmn falls between λ and $\lambda + d\lambda$. If the photon energy measured from the Franck–Condon excitation energy is normalized by

$$x = (E - E_a)/(2c'^2 k_B T)^{1/2}, \quad (4.7.4)$$

the corresponding lineshape function $f(x)$, defined by $F(E)dE = f(x) dx$, can be written as

$$f(x) = 3^{-1} \pi^{-3/2} \sum_{i=1}^3 \int_{-\gamma}^{\gamma} x^2 t_i(\lambda)^{-3} \exp[-x^2 t_i(\lambda)^{-2}] (d\Omega/d\lambda) d\lambda. \quad (4.7.5)$$

The differential solid angle is calculated as

$$(d\Omega/d\lambda) = \int_0^\pi \int_0^{2\pi} \sin \theta d\theta d\phi \delta[(1/2)(\cos \theta \sin^2 \theta \sin 2\phi) - \lambda]. \quad (4.7.6)$$

As is readily seen, eq. (4.7.6) diverges logarithmically as $|\lambda| \rightarrow 0$. The results of numerical computation of (4.7.6) and (4.7.5) are shown in Figs. 4.6 (b) and 4.7. The logarithmic divergence in the former is brought into the latter. What is interesting is the split lineshape caused by the split adiabatic potential in spite of the directional average. Since $d\Omega/d\lambda$ diverges as $|\gamma| \rightarrow 0$, an important contribution to the lineshape (4.7.5) comes from the region with small $|\lambda|$ with well-split $t_i(\lambda)$ s in Fig. 4.6(a). The dispersion D of this lineshape is given by $D^2 = (1/3)\text{Trace}\langle(H_1)^2\rangle = (3/2)c'^2 k_B T$.

To sum up, we have seen three different types of interaction mode as regard their roles in *removal of degeneracy* of excited states dipole-allowed from the totally symmetric ground state and in *splitting of the* corresponding *absorption band* (in the adiabatic limit in which vibrational quanta are negligible). The first is the *inactive mode* which causes *no level splitting* and hence *no band splitting* (totally symmetric A_{1g}). The second is the *potentially active mode* which gives rise to *level splitting but not to band splitting* (two-dimensional E_g). The third is the *active mode* which gives rise to *level splitting and also to band splitting* (three-dimensional T_{2g}). This classification of interaction modes is extended to systems with various symmetries and tabulated in Table 4.1.³ There are only two possible types of split lineshape. One is the triplet lineshape already shown in Fig. 4.7. The second type with a doublet structure appears with the E mode in the electronic transition $A \rightarrow E$ in a C_3 system,

Table 4.1 Active and potentially active modes of lattice vibrations around the point imperfection for a dipole-allowed electronic transition from the totally symmetric ground state to a degenerate excited state.

Symmetry of the point imperfection	Electronic transition	Active modes	Potentially active modes
$C_4 = 4$ $S_4 = \bar{4}$ $C_{4h} = 4/m$ $C_{4v} = 4 \text{ mm}$ $D_{2d}(V_d) = \bar{4}2 \text{ m}$ $D_4 = 422$ $D_{4h} = 4/mmm$	$A \rightarrow E$ $A_g \rightarrow E_u$ $A_1 \rightarrow E$ $A_{1g} \rightarrow E_u$		B B_g B_1, B_2 B_{1g}, B_{2g}
$C_3 = 3$ $S_6(C_{3i}) = \bar{3}$ $C_{3v} = 3 \text{ m}$ $D_3 = 32$ $D_{3d} = \bar{3} \text{ m}$ $C_6 = 6$ $C_{3h} = \bar{6}$ $C_{6h} = 6/m$ $C_{6v} = 6 \text{ mm}$ $D_6 = 622$ $D_{3h} = \bar{6} \text{ m}2$ $D_{6h} = 6/mmm$	$A \rightarrow E$ $A_g \rightarrow E_u$ $A_1 \rightarrow E$ $A_{1g} \rightarrow E_u$ $A \rightarrow E_1$ $A' \rightarrow E'$ $A_g \rightarrow E_{1u}$ $A_1 \rightarrow E_1$ $A'_1 \rightarrow E'$ $A_{1g} \rightarrow E_{1u}$	E E_g E E_g E_2 E' E_{2g} E_2 E' E_{2g}	
$T = 23$ $T_d = \bar{4}3 \text{ m}$ $T_h = m 3$ $O = 432$ $O_h = m 3 \text{ m}$	$A \rightarrow T$ $A_1 \rightarrow T_2$ $A_g \rightarrow T_u$ $A_1 \rightarrow T_1$ $A_{1g} \rightarrow T_{1u}$	T T_2 T_g T_2 T_{2g}	E E E_g E E_g

Source: Toyozawa, Y. and Inoue, M. (1996) *J. Phys. Soc. Jpn* **21**, 1663.

representatively. The adiabatic potentials and the absorption lineshapes are shown in Figs. 4.8 and 4.9, respectively. Quantum theoretical treatment of this system has been performed by Longuet-Higgins *et al.*¹² Figure 4.9 almost coincides with the envelope of the densely distributed line spectra they calculated in the situation that the vibrational quantum is small compared with the classical band width $(a^2 k_B T)^{1/2}$.

Throughout the study of the various systems shown in Table 4.1, one realizes that the origin of the difference between the potentially active and the active mode is whether the Franck–Condon point is merely an intersection point of the adiabatic potentials or a branching point.

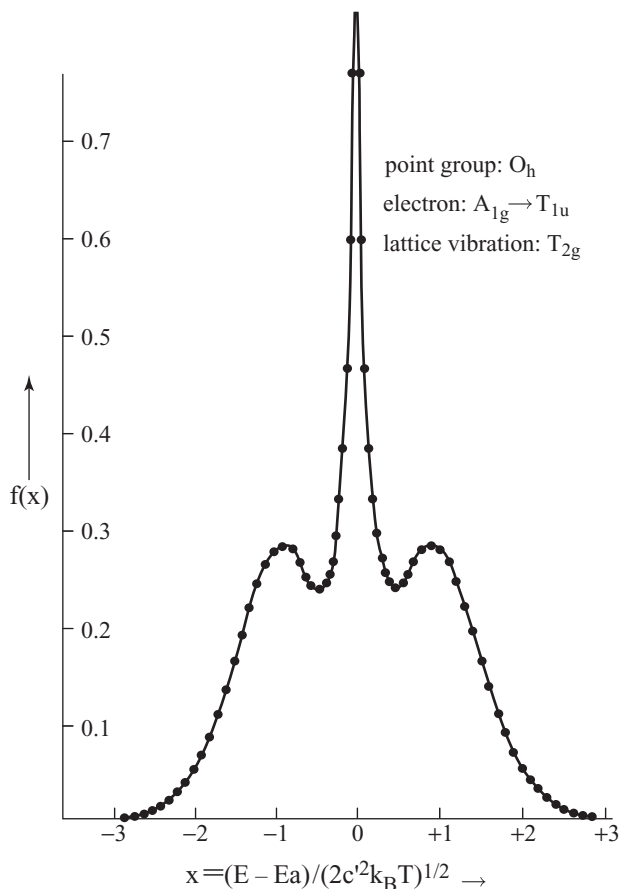


Fig. 4.7 Calculated absorption band for the optical transition $A_{1g} \rightarrow T_{1u}$ of a localized electron at a defect with O_h symmetry, with Jahn-Teller splitting through the interaction of the T_{1u} state electron with T_{2g} modes of the lattice vibration.³

Extensive study of the lineshapes of the $A_{1g} \rightarrow T_{1u}$ transition in the O_h system has been performed by Cho¹³ for the most general case of nonvanishing coupling coefficients a , b and c of the interactions with A_{1g} , E_g and T_{2g} modes, respectively. In Fig. 4.10 are reproduced only the results for $a = 0$, $b \geq 0$ and $c > 0$ since inclusion of $a \neq 0$ results in smearing of the splitting through the convolution of the width of the Gaussian due to a alone and may be conjectured qualitatively from this figure. One can see how the mixing of the potentially active mode E_g with the active mode T_{2g} changes the lineshape. Close observation reveals that the height of the peaks on both sides becomes comparable to that of the central peak in the intermediate cases around $b/c \sim 1$, while the central peak becomes a small hump on the broader band (bell shape) at $4 < b/c (=b'/c') < 9$, tending to a simple Gaussian as $b/c \rightarrow \infty$. This is a characteristic feature of mixing the

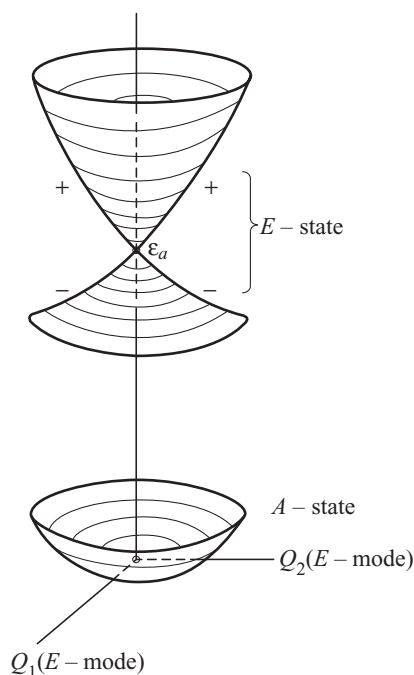


Fig. 4.8 Adiabatic potential surfaces for the electronic states A and E of a system with trigonal symmetry C_3 , in the two-dimensional configuration coordinate space (Q_1 , Q_2) of the E mode.³

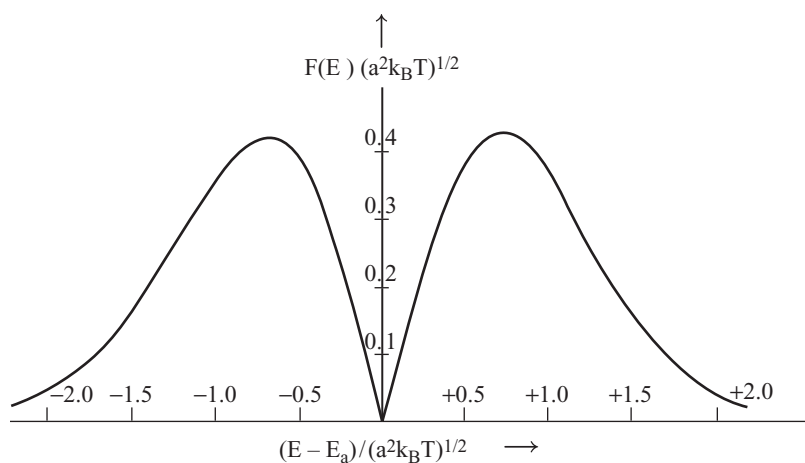


Fig. 4.9 Calculated absorption band for the optical transition $A \rightarrow E$ in Fig. 4.8, reflecting the Jahn-Teller-split adiabatic potential of the E state.³

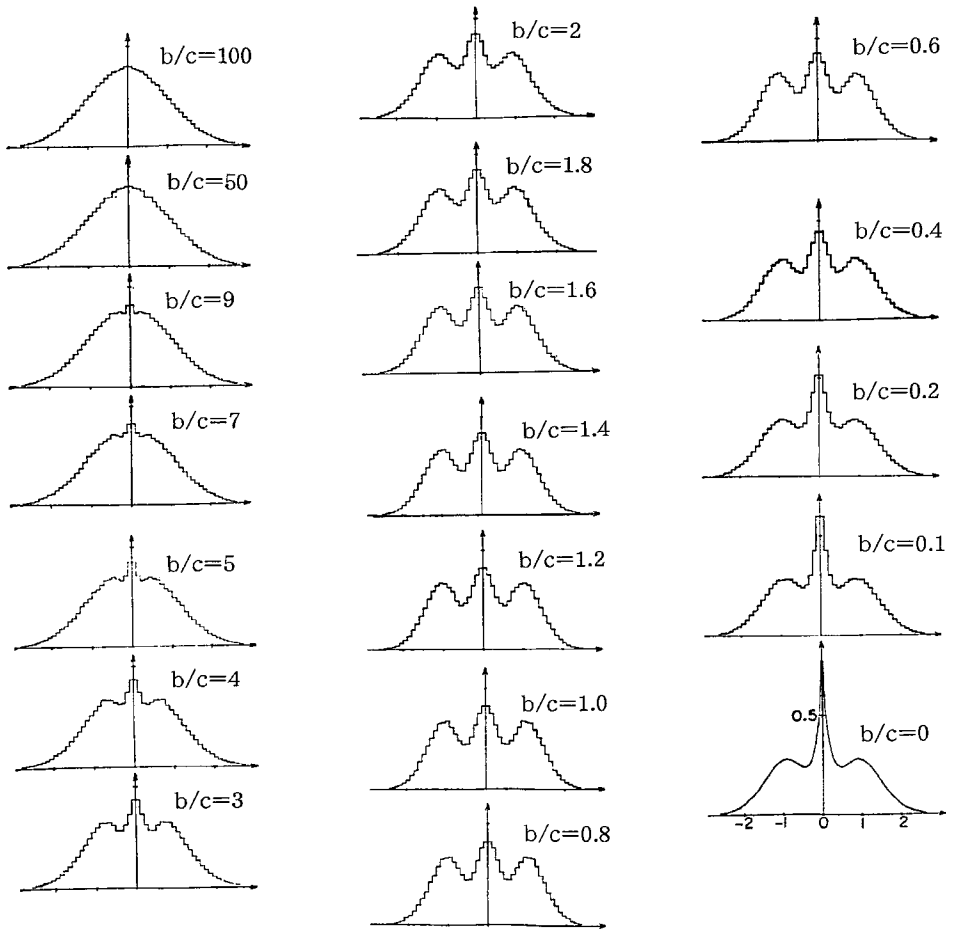


Fig. 4.10 Calculated lineshape of an electronic system, such as that in Fig. 4.7, interacting with E_g and T_{2g} modes with various ratios b/c ($= b'/c'$) of coupling coefficients. The abscissa is normalized in units of the square root of the second moment of the lineshape which is proportional to $\{k_B T [b^2 + (3/2)c^2]\}^{1/2}$. Due to Cho.¹³

potentially active mode which can never be seen in the case of mixing the inactive mode.

We will now discuss the experimental studies which motivated the theoretical studies mentioned above. Yuster and Delbecq¹⁴ observed the triplet structure in the C band of KI: Tl, which they suspected to be due to the Jahn-Teller effect of the excited state. Subsequently, several authors^{15,16} have observed the structures in the A and C bands of various alkali halide phosphors. Systematic studies of the absorption spectra, for various combinations of host alkali halides and impurity ions, have been carried out by Fukuda *et al.*¹⁷ They revealed that, in contrast to the triplet structure of the C band (see Fig. 4.11), the A band has doublet structures (see

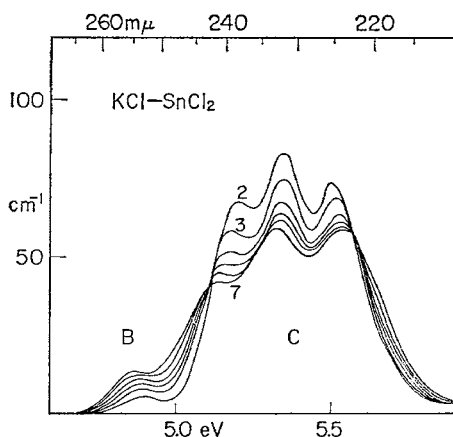


Fig. 4.11 Observed B and C absorption bands in KCl:Sn at various temperatures: 2) 77 K, 3) 133 K, 4) 175 K, 5) 211 K, 6) 243 K, 7) 285 K. Due to Fukuda *et al.*¹⁷

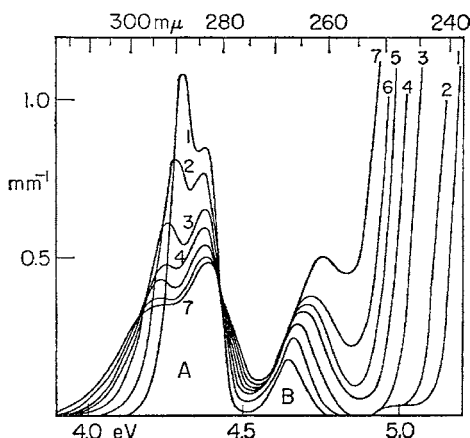


Fig. 4.12 Observed A and B bands in KCl:In at various temperatures: 1) 20 K, 2) 77 K, 3) 133 K, 4) 175 K, 5) 211 K, 6) 243 K, 7) 285 K. Due to Fukuda *et al.*¹⁷

Fig. 4.12). The splitting energies of these bands increase as temperature rises, which fact strongly supports vibration-induced splitting as the underlying mechanism. The possible participation of a great number of normal modes made quantitative theoretical description of these spectral structures difficult.

In view of the observed temperature-dependent splitting *without* phonon structures, the concept of the interaction mode was introduced³ which reduced the N -dimensional C.C. space to a six-dimensional one, thereby facilitating quantitative study of lineshapes though within the adiabatic limit (namely, at the cost of replacing the quantal description of lattice vibrations by the classical one). As a result, the

triplet structure of the C band has been explained at a quantitative level.¹³ In fact, the observed lineshapes of the C band in various combinations of guest *monovalent* metal ions (see below for *divalent* ions) and host alkali halides can be almost fitted to some of the lineshapes in Fig. 4.10, except for the asymmetry in the former which is presumably due to negative curvatures of the adiabatic potentials originating from interaction with higher energy excited states.

In order to explain the peculiar lineshape of the A band, one must take account of the exchange and spin-orbit interactions in the excited states with (ns) (np) configurations which result, in order of *decreasing* energy, in spin-singlet-originated 1P_1 ($^1T_{1u}$) states {C band}, and triplet-originated 3P_2 ($^3E_u + ^3T_{2u}$) states {B band}, 3P_1 ($^3T_{1u}$) states {A band} and 3P_0 ($^3A_{1u}$) states {not observed}. Here the notation within a parenthesis (\cdots) is exact for cubic symmetry (O_h) while that before the parenthesis is the corresponding state within the atomic (spherical) approximation. Within the curly brackets $\{\cdots\}$ are the observed absorption band assigned by Seitz¹⁸ to correspond to the excited states mentioned there. The last three of the four states mentioned above are separated from the first one by the exchange energy and from each other by the spin-orbit interaction which also has the effect of mixing the first one with the third one, making the latter optically allowed. If one *assumes* that the matrix elements of the electron-phonon interaction H_I between the first and the third and within the third states consist of the same interaction modes as were introduced within the first states, one can calculate the lineshape of the A band using the same interaction coefficients, the observed energy difference and the intensity ratio of the C and A bands (governed by the exchange and spin-orbit interaction) as discussed by Sugano¹⁹ in his extension of Seitz's argument to the molecular orbital model. The result of a calculation with $c \neq 0$, $a = b = 0$ is shown in Fig. 4.13. The asymmetry, with the lower-energy tail growing with temperature, is due to the negative curvature of the adiabatic potential originating from interaction with the $^1T_{1u}$ state through H_I – the pseudo Jahn-Teller effect. It reproduces the observed doublet structure with temperature-dependent splitting and asymmetry (inclusive of reversal of relative peak heights) shown in Fig. 4.12.

The symmetry of the relaxed excited states (static Jahn-Teller effect) has also been studied with the use of polarized light at the A absorption band and by detecting the polarization of emitted light, in correlation with that of the incident light and as referred to the crystal axis. The correlation turned out to be greater for the incident light at the lower side of the absorption peak than at the higher side, as is expected from the Jahn-Teller split adiabatic potential.²⁰ The study of the polarization of emitted light under uniaxial stress and isotropic pressure by Shimada and Ishiguro²¹ and by Masunaga *et al.*²² revealed the fine detail of the nature and symmetries of the relaxed excited states which can be explained only by considering the Q_1 -dependences of the coupling coefficients b and c .

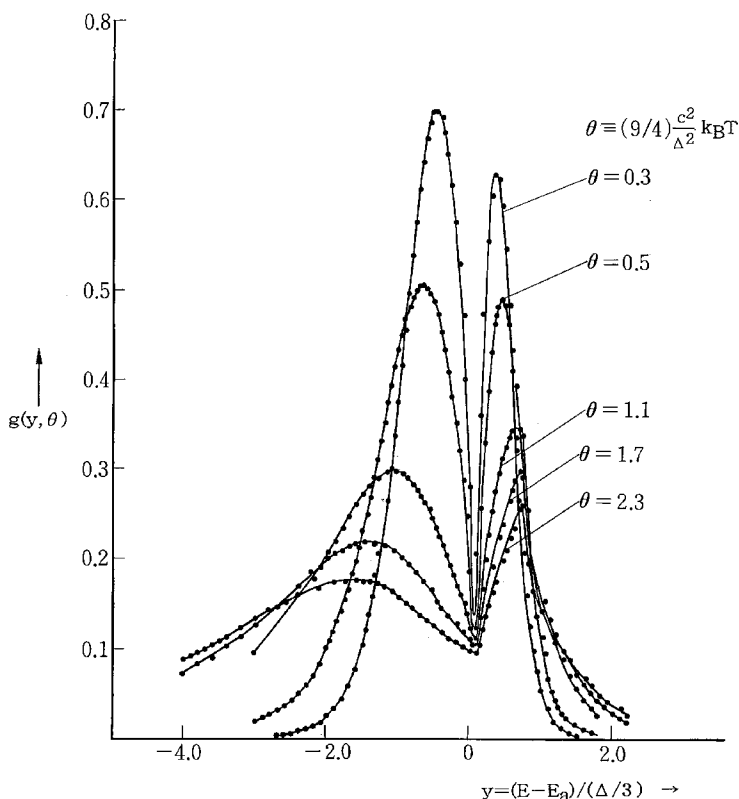


Fig. 4.13 The calculated lineshapes of the A band at various temperatures in the limit of no triplet-singlet mixing. Note the temperature-dependent asymmetry and intensity ratio of the split peaks which are due to the pseudo Jahn-Teller effect of being repelled from the C band (situated above the A band) through the T_{2g} mode interaction, the only mode considered here.³

The coupling coefficients a , b and c can be estimated with the quasi-molecular model in which the central metal ion and neighboring six halide ions, with their Coulomb interactions and the electronic polarization of the latter, are considered and use is made of suitably corrected *local* elastic moduli instead of the individual interatomic forces. Reasonable agreements with observations are obtained, inclusive of the general tendency that divalent ions have greater coupling coefficients than monovalent ones because of the polarization due to the extra charge. A certain part of the discrepancy between the observed lineshapes from those of Fig. 4.10 in the case of *divalent* ions, however, are supposed to come from an associated alkali-ion vacancy on the second-nearest site which neutralizes the *extra* charge, a static origin of the splitting.²³

Recently, Ueda *et al.*²⁴ measured the absorption band due to the carbon 1s electron of CF_4 molecules using high-resolution X-ray spectroscopy. The results were

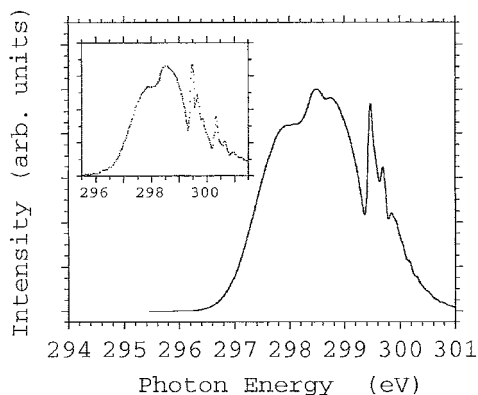


Fig. 4.14 Quantum-mechanically calculated C 1s absorption band of CF_4 molecule. The parameter values are taken so as to reproduce the experimental results (shown in the inset). Due to Itoh *et al.*²⁵

analysed theoretically by Itoh, Tanaka and Kayanuma.²⁵ The final state of the optical transition is a T_2 state composed of four C–F antibonding σ^* orbitals, and this main absorption band is followed on its higher-energy side by a series of Rydberg states. The former is subject to (1) the Jahn-Teller effect due the T_2 mode vibration composed of four C–F stretching modes, (2) the pseudo Jahn-Teller effect due to the interaction with the optically forbidden A state a little above it and the aforementioned Rydberg states through the same T_2 mode, and (3) the interaction with the totally symmetric S mode vibration consisting of the stretching modes, being a system somewhat analogous to the Ti^+ -like center in an alkali-halide crystal. Due to the simplicity of the molecule, the T_2 interaction mode is exactly a normal mode, there being no other normal modes with the same symmetry. This facilitated a full quantum mechanical calculation of the absorption spectrum, the result of which is in fact in excellent agreement with the observation down to the fine detail of the asymmetrically split T_2 band and the Fano antiresonance (see Sections 10.2 and 12.7) between the T_2 band and the Rydberg states, as shown schematically in Fig. 4.14 and its inset (observed data). The result was also compared with the semi-classical calculation, revealing the relative contributions to the asymmetry of the quantum effect and the pseudo Jahn-Teller effect. Experimentally, photo-excitation to the T_2 state results in dissociation of the F atoms, indicating that in the relaxed excited state of the molecule one of the four C–F bonds is infinitely stretched.

4.8 Phonon coupling strength of localized electrons in solids

In contrast to the Ti^+ -like center in which the relevant electrons are well localized within the Ti^+ ion, the electron in the F center is spread over several and tens of nearby host ions in its ground and excited states, respectively. Localized electrons

in host crystals with smaller energy gaps and hence with larger dielectric constants, such as donor electrons in Si, Ge and III–IV semiconductors, have orbital radii as large as several to ten times the lattice constant. In eq. (4.5.12) we defined the coupling strength, S , characterizing the total width of the optical absorption band. This width reduces to $D \sim S^{1/2} \langle \hbar \omega \rangle$ at the absolute zero of temperature, see eq. (4.5.17). In order to see the dependence of S on the orbital radius of the electron, which varies as much as 10^2 times depending on the center as mentioned, let us consider a dipole-allowed transition from an s-like ground state to a p_x -like excited state with wave functions given by

$$\phi_g = (\alpha^3/\pi)^{1/2} \exp(-\alpha r), \quad \phi_e = (\beta^5/\pi)^{1/2} x \exp(-\beta r) \quad (4.8.1)$$

with respective mean square radii

$$r_g^2 \equiv \langle r^2 \rangle_g = 3\alpha^{-2}, \quad r_e^2 \equiv \langle r^2 \rangle_e = (15/2)\beta^{-2}. \quad (4.8.2)$$

We replace the normal modes q_i of the lattice vibrations, ignoring the perturbation of the lattice defect, by those of the host lattice which consist of acoustic and optical modes as is typical in ionic crystals. Namely, we do not distinguish the modes of different symmetries around the defect such as A_{1g} , E_g and T_{2g} , but consider them *altogether*. The q_i -dependent part of the potential $V(\mathbf{r}, q)$ for the electron is expanded as powers of $\{q_i\}$, of which the linear term can be written as

$$H_I = \sum_i V_i(r) q_i \quad (4.8.3)$$

with the coefficients being given by

$$V_i(\mathbf{r}) \equiv V_{\mathbf{m}, \mathbf{q}, \sigma}(\mathbf{r}) = V_{\mathbf{m}, \mathbf{q}} \times \begin{cases} \cos(\mathbf{q} \cdot \mathbf{r}) & (\sigma = 1), \\ \sin(\mathbf{q} \cdot \mathbf{r}) & (\sigma = 2), \end{cases} \quad (4.8.4)$$

$$V_{\mathbf{m}, \mathbf{q}} = \begin{cases} [2\Xi^2/NM]^{1/2} q & (m = \text{acoustic mode}) \\ [(\epsilon_e^{-1} - \epsilon_s^{-1})2e^2\omega_\ell^2/Nv_0] q^{-1} & (m = \text{optical mode}) \end{cases} \quad (4.8.5)$$

where N is the total number of unit cells in the crystal, v_0 and M the volume and mass, respectively, of a unit cell, Ξ the deformation potential constant, ω_ℓ the angular frequency of the longitudinal optical mode, ϵ_e and ϵ_s the electronic and static dielectric constants, respectively. The wave vector \mathbf{q} with its magnitude $q \equiv |\mathbf{q}|$ should not be confused with the normal mode coordinate q_i . We leave the derivation of (4.8.4, 4.8.5) to later sections (see eqs. (9.1.15, 9.1.16, 9.1.18) and (9.4.3)). Since we make use of the normal mode coordinates q_i (real quantities) for the standing waves $\cos(\mathbf{q} \cdot \mathbf{r})$ and $\sin(\mathbf{q} \cdot \mathbf{r})$ in place of the annihilation–creation operators b and b^\dagger (see (2.2.5)) of phonons for running waves $\exp(i\mathbf{q} \cdot \mathbf{r})$ in Chapter 9, we have to confine the \mathbf{q} -summation to half of the \mathbf{q} -space (say to $q_z > 0$) here.

With the use of (4.3.9) and (4.4.8), the displacement of the normal coordinate q_i due to the optical transition and the coupling function are given, respectively, by

$$\Delta_i = -(V_{iee} - V_{igg})/\omega_i^2, \quad (4.8.6)$$

$$s(\omega) = \sum_i (\omega_i \Delta_i^2 / 2\hbar) \delta(\omega_i - \omega). \quad (4.8.7)$$

In the case of an acoustic mode with the dispersion relation $\omega_{ac,q} = uq$ ($q \rightarrow 0$), where u is the longitudinal sound velocity, one readily finds that $\Delta_{ac,q} \propto q^3$ as $q \rightarrow 0$ and $s(\omega) \propto \omega^5$ as $\omega \rightarrow 0$. The expressions for $s(\omega)$ and $S \equiv \int d\omega s(\omega)$ are somewhat lengthy. However, if one assumes that the orbital radius of the excited state is significantly larger than that of the ground state ($r_p > 2r_s$), one obtains the following approximate expressions in both the limits of a small- and large-radius electron:

$$(I) \quad r_p < v_0^{1/3} :$$

$$S_{ac} \sim 0.46[\Xi^2/\hbar M u^3](r_p/v_0^{1/3})^4, \quad (4.8.8)$$

$$S_{op} \sim 1.8[(\epsilon_e^{-1} - \epsilon_s^{-1})e^2/4\pi\hbar\omega_\ell v_0^{1/3}](r_p/v_0^{1/3})^4, \quad (4.8.9)$$

$$(II) \quad v_0^{1/3} < r_s$$

$$S_{ac} \sim 0.051[\Xi^2/\hbar M u^3](r_s/v_0^{1/3})^{-2}, \quad (4.8.10)$$

$$S_{op} \sim 0.54[(\epsilon_e^{-1} - \epsilon_s^{-1})e^2/4\pi\hbar\omega_\ell v_0^{1/3}](r_s/v_0^{1/3})^{-1}. \quad (4.8.11)$$

It is interesting to note that the electron radius closer to the lattice constant (r_p in case (I) and r_s in case (II)) is determinative of the coupling strength, namely the electron interacts most effectively when its radius is comparable to the lattice constant. This is also seen in the fact that both S_{ac} and S_{op} given above become largest when the orbital radius approaches the lattice constant. Putting values of parameters typical for ionic crystals, the S s are plotted as functions of the determinative orbital radius in Fig. 4.15.⁴ The different powers in the r dependence in case (II) come from the fact that the electron–phonon interaction is of short range in the case of the deformation potential (see Section 9.4) and of long range in the case of electrostatic interaction (see Section 9.1).

As for the coupling function $s(\omega)$, one should note an important difference between cases (I) and (II) for the acoustic mode interaction. While the small-radius electron couples indiscriminately with phonons of any wavelength, the large-radius electron interacts most effectively with phonons of wavelength comparable to the radius and $s(\omega)$ decays rapidly when ω exceeds the corresponding frequency. On the other hand, in the low-frequency region the factor ω^5 appears in both cases (I) and (II), the coefficient being proportional to the fourth power of r_p , the larger

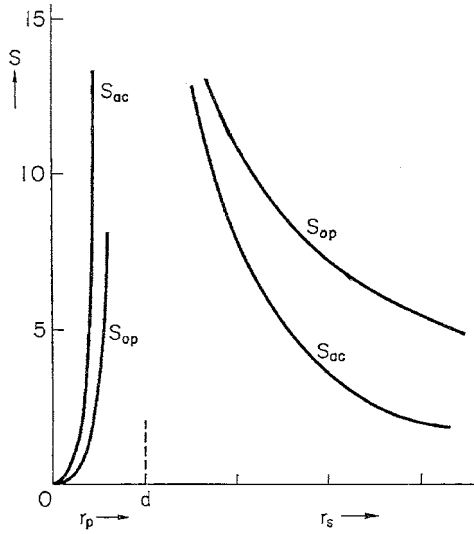


Fig. 4.15 Coupling strength S for acoustic and optical modes in a typical ionic crystal, calculated as functions of the orbital radius r of a localized electron following eqs. (4.8.8) to (4.8.11). The parameters used are: $v_0 = 0.7 \times 10^{-22} \text{ cm}^3$, $M/v_0 = 3 \text{ g cm}^{-3}$, $\Xi = 6 \text{ eV}$, $c_\ell = 0.3 \times 10^6 \text{ s}^{-1}$, $\hbar\omega_{\text{op}} = 0.015 \text{ eV}$, $\epsilon_e/\epsilon_0 = 3$, $\epsilon_s/\epsilon_0 = 6$.⁴

of the electron radii which is more effective in interacting with long-wavelength phonons. The two cases, with equal coupling strength S_{ac} , are shown schematically in the left half of Fig. 4.16. In the case of the optical mode interaction, the large-radius electron interacts most effectively with long-wavelength phonons because of the long-range interaction (see the factor q^{-1} in (4.8.5)), in contrast to the case of the small-radius electron where $s(\omega)$ extends over the entire range of the optical phonon frequency $\omega_{\text{op}}(\mathbf{k})$ as shown in the right half of Fig. 4.16.

From Figs. 4.15 and 4.16, one can find favorable and unfavorable situations for the observation of phonon structures as follows. For a localized electron with orbital radius comparable to the lattice constant, the coupling strength S is too large (several tens) for the phonon structure to be observable. The F center and the TI^+ -like center in alkali halides belong to this type. In fact, from the observed width or splitting one can estimate the dispersion D and thence the coupling strength S , making use of (4.5.17) and a suitable value of the average phonon frequency, to obtain $S \sim 20\text{--}40$, in qualitative agreement with $S = S_{\text{ac}} + S_{\text{op}}$ in Fig. 4.15. The zero-phonon line, with expected fractional intensity $\exp(-S)$, is too small to be observed.

An electron radius moderately large compared to the lattice constant is favorable for the appearance of a number of equidistant optical phonon lines as sidebands, not only because of the moderate value of $S_{\text{op}}(>S_{\text{ac}})$ but also because of the sharp peak in the spectral coupling $s(\omega)$ as seen in Fig. 4.16. The well-known edge emission

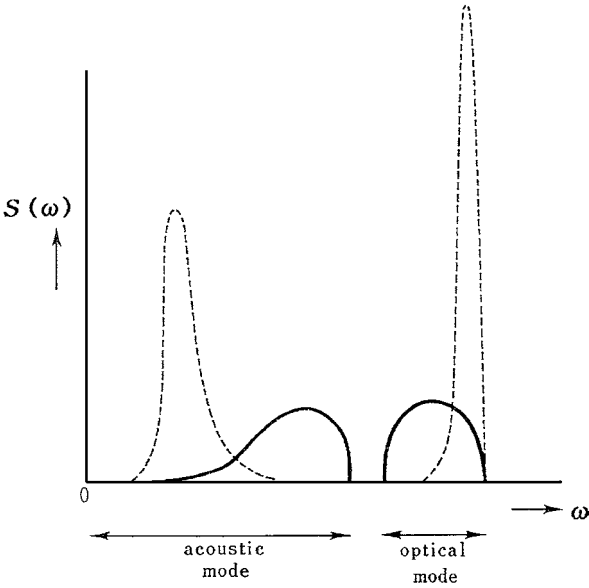


Fig. 4.16 Coupling function $s(\omega)$ for the cases of (I) small radius (solid line) and (II) large radius (dashed line) of a localized electron.⁴

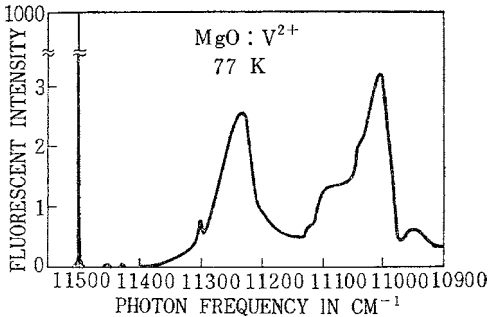


Fig. 4.17 Phonon structure associated with 8697 Å fluorescence line of $\text{MgO}:\text{V}^{2+}$. Due to Sturge *et al.*²⁷

band in CdS, and the mirror symmetric set of absorption and fluorescence bands in ZnTe seem to belong to this case.²⁶ In the latter example, each of the equidistant multiple optical phonon lines is accompanied by a weak acoustic one-phonon sideband with the shape $s_{\text{ac}}(\omega)$.

An electron well localized within the impurity atom is more favorable for the observation of the coupling function $s(\omega)$ separately since multiphonon sidebands are expected to be weak. As a typical example, the fluorescence spectrum of V^{2+} in MgO is shown in Fig. 4.17.²⁷ Here one finds that the acoustic phonon sideband is as strong as the optical one, in agreement with Fig. 4.15.

5

Lattice vibrations

5.1 Simple crystal lattices

Atoms in molecules and solids are bound together by the valence electrons which have been supplied by them and are *shared* among them. In this chapter we will consider the motion of atoms under *given* interatomic forces *without* asking about their nature – the electronic origin of the bonds. We will confine ourselves to solids¹ and study some simple characteristics of the atomic motions within them – lattice vibrations² – brought about by the periodic structure of the crystal lattice. The lattice vibrations will turn out to be a prototype of *elementary excitations* in a solid, other examples of which will be given in later sections.

Consider an assemblage of a great number (N) of identical atoms with interatomic potential $v(r)$ which typically is attractive except for the hard-core repulsion, and let them be arrayed in one-dimensional space with positions x_1, x_2, \dots in increasing order. In the approximation that only the nearest neighbor potentials are considered, as is allowed when $v(r)$ is short ranged, the total potential energy is given by $U \simeq v(x_1 - x_2) + v(x_2 - x_3) + \dots$, and the most stable arrangement of the atoms is a periodic array with equal distance a at which $v(r)$ is a minimum. If the atoms are slightly displaced by ξ_n from their equilibrium positions $x_n = na$ ($n = 1, 2, \dots$), they are subject to the restoring forces derivable from the potential energy which is Taylor-expanded up to the quadratic terms (harmonic approximation)

$$U = (N - 1)v(a) + [v''(a)/2][(\xi_1 - \xi_2)^2 + (\xi_2 - \xi_3)^2 + \dots]. \quad (5.1.1)$$

The first term with reversed sign represents the binding energy of this one-dimensional crystal but for the vibrational motion of the atoms. The equations of motion are given by

$$M d^2 \xi_n / dt^2 = -\partial U / \partial \xi_n = f[(\xi_{n+1} - \xi_n) - (\xi_n - \xi_{n-1})], \quad (5.1.2)$$

where M is the atomic mass and $f \equiv v''(a)$ is the force constant.

Let us study the solution of (5.1.2) in which all atoms oscillate with the same angular frequency ω , the so-called *normal mode*. Putting

$$\xi_n(t) = A_n \exp(-i\omega t) \quad (5.1.3)$$

in (5.1.2), one obtains simultaneous equations for the amplitudes

$$-M\omega^2 A_n = f(A_{n+1} - A_n) - f(A_n - A_{n-1}). \quad (5.1.4)$$

Being of identical form for all n , any solution of (5.1.4) should be such that successive A_n s have a common ratio γ , namely $A_n = A\gamma^n$. In order to avoid the situation of infinitely increasing or decreasing amplitude towards the ends, one should take $|\gamma| = 1$, namely one can take

$$A_n = AN^{-1/2} \exp(ikna), \quad (5.1.5)$$

where k can take *any real* number. The situation is made even clearer if one considers a chain of $N(\gg 1)$ atoms with finite but macroscopic length $L = Na$ and imposes a *cyclic boundary condition* (which amounts to imagining a ring of the same length)

$$\xi_{N+1} = \xi_1, \quad (5.1.6)$$

as was done in Section 1.1 for electromagnetic waves. Then one finds that k can take only *discrete* values given by

$$k = (\ell/N)b \quad (5.1.7)$$

with integer ℓ , where b is a *basic reciprocal lattice* defined by

$$b \equiv 2\pi/a \quad (5.1.8)$$

(compare with eq. (1.3.6)).

In contrast to the continuous space considered in Section 1.3, however, one should note that any two wave vectors k and k' give identical waves when their difference is an integer times b since (5.1.5) is a periodic function of k with period b . In order to have mutually independent solutions, one has to confine k to an interval of finite length b , corresponding to confining ℓ to N successive integer values according to (5.1.7) as is consistent with the number of degrees of freedom.

Putting (5.1.5) into (5.1.4), one readily obtains the dispersion relation between ω and k :

$$\omega_k = 2(f/M)^{1/2} |\sin(ka/2)| \quad (5.1.9)$$

where only a positive root of ω has been taken by definition. For $N \gg 1$, k given by (5.1.7) is an almost continuous variable, and (5.1.9) is a continuous function of k as shown in Fig. 4.2. We will confine k to the interval $[-b/2, +b/2]$, a symmetric choice among other possible intervals with length b .

The general solution of (5.1.2) is given by a linear combination of the normal modes given by (5.1.3) and (5.1.5) with (5.1.7), which we write as

$$\begin{aligned}\xi_n(t) &= (2NM)^{-1/2} \sum_k [A_k \exp(ikna - i\omega_k t) + (c.c.)] \\ &= (2NM)^{-1/2} \sum_k [\eta_k(t) + \eta_{-k}^*(t)] \exp(ikna).\end{aligned}\quad (5.1.10)$$

Here (c.c.), indicating the conjugate complex of the preceding term, has been added to assure that $\xi_n(t)$ is real *irrespective* of the choice of complex numbers A_k and, after rewriting k by $-k$, has been united in the second line with the original term $\eta_k(t) \equiv A_k \exp(-i\omega_k t)$ (which is needed only for the purpose of expanding $\xi_n(t)$ in the complete set: $\exp(ikna)$). Thus one can take any initial values of N positions and N momenta by choosing N complex numbers A_k appropriately.

With the use of (5.1.10), the equation of motion: $d\eta_k(t)/dt = -i\omega_k \eta_k(t)$ and the orthogonality relation

$$N^{-1} \sum_n \exp[i(k - k')na] = \delta_{k,k'}, \quad (5.1.11)$$

the kinetic energy can be written as

$$K = (M/2) \sum_n [d\xi_n(t)/dt]^2 = (1/4) \sum_k (\omega_k)^2 (\eta_k - \eta_{-k}^*)(\eta_k^* - \eta_{-k}).$$

Combining this with the similar expression for U of eq. (5.1.1), one obtains the total energy

$$H = K + U = (1/2) \sum_k \omega_k^2 \eta_k \eta_k^* = (1/2) \sum_k (P_k^2 + \omega_k^2 Q_k^2) \quad (5.1.12)$$

with the use of (5.1.9) and the real variables Q_k and P_k defined by

$$\eta_k = (Q_k + iP_k/\omega_k)/\sqrt{2}. \quad (5.1.13)$$

Interpreting Q_k and P_k as the new coordinate and momentum, respectively, (though both of them are mixtures of the original coordinate η_k and momentum $d\eta_k/dt$), one obtains a harmonic oscillation as the solution of the equation of motion with the Hamiltonian H .

For long wavelengths with k much smaller than the basic reciprocal lattice b , $\xi_n(t)$ given by (5.1.3) and (5.1.5) is a slowly varying function of n and can be considered as a continuous function of a variable $x \sim x_n = na$. Then one can approximate $(\xi_{n+1} - \xi_n)/a$ by a derivative $d\xi(x)/dx$ at $x = (n + \frac{1}{2})a$, the quantity in $[\dots]$ of eq. (5.1.2) by a^2 times the second derivative $d^2\xi/dx^2$ at $x = na$, and the difference equation (5.1.2) by a differential equation

$$\partial^2 \xi / \partial t^2 = (C/\rho) \partial^2 \xi / \partial x^2. \quad (5.1.14)$$

Here $\rho \equiv M/a$ is the density and $C \equiv fa$ represents an elastic compliance defined as the (tensile) stress $\sigma_T = f(\xi_{n+1} - \xi_n)$ divided by the strain $(\xi_{n+1} - \xi_n)/a$. Both ρ

and C are *macroscopic* quantities, and the equation of motion (5.1.14) for an elastic *continuum* has a *sound-wave* solution $\xi(x, t) \propto \exp(ikx - i\omega t)$ with *sound velocity* $c_s \equiv \omega/k$ being given by $(C/\rho)^{1/2}$ in agreement with the long-wavelength limit of (5.1.9). For this reason, the solution (5.1.3) with (5.1.5) is called the *acoustic* mode of lattice vibration, in contradistinction to the *optical* modes with non-vanishing ω_k at $k = 0$ to be mentioned later.

A real crystal consists of a three-dimensional lattice with periodicity specified by three *basic lattice vectors* $\mathbf{a}_1, \mathbf{a}_2$ and \mathbf{a}_3 which are not coplanar but not necessarily orthogonal to each other, such that the translation from any point in the crystal by a *lattice vector*

$$\mathbf{R}_{n_1, n_2, n_3} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (5.1.15)$$

with arbitrary integers n_i brings one to an equivalent environment. A unit cell, a parallelepiped spanned by $\mathbf{a}_1, \mathbf{a}_2$ and \mathbf{a}_3 , has volume $v_0 = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)$ where \cdot and \times denote scalar and vector products, respectively.

Let us confine ourselves for the moment to simple crystal lattices of which a unit cell contains one atom, deferring the description of more general types of crystal to the subsequent section. One can then assume that the atoms are arrayed on the lattice points $\mathbf{R}_{n_1, n_2, n_3}$ given by (5.1.15).

Similarly to the one-dimensional lattice, one can readily find that the normal modes of the lattice vibrations in a three-dimensional lattice consist of running waves:

$$\xi_{n_1, n_2, n_3}(t) = \xi \exp(ik \cdot \mathbf{R}_{n_1, n_2, n_3} - i\omega t). \quad (5.1.16)$$

To see the allowed values of \mathbf{k} , it is convenient to define three *basic reciprocal lattice vectors* $\mathbf{b}_1, \mathbf{b}_2$ and \mathbf{b}_3 in the \mathbf{k} -space by

$$\mathbf{b}_1 = (2\pi/v_0) \mathbf{a}_2 \times \mathbf{a}_3, \quad \mathbf{b}_2 = (2\pi/v_0) \mathbf{a}_3 \times \mathbf{a}_1, \quad \mathbf{b}_3 = (2\pi/v_0) \mathbf{a}_1 \times \mathbf{a}_2, \quad (5.1.17)$$

and the *reciprocal lattice vectors* by

$$\mathbf{K} = K_1 \mathbf{b}_1 + K_2 \mathbf{b}_2 + K_3 \mathbf{b}_3 \quad (5.1.18)$$

with arbitrary integers K_i . The basis vectors in \mathbf{r} - and \mathbf{k} -spaces satisfy the relations

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi \delta_{ij}. \quad (5.1.19)$$

Let us first note that the waves (5.1.16) for \mathbf{k} and \mathbf{k}' are equivalent when $\mathbf{k} - \mathbf{k}'$ is a reciprocal lattice vector \mathbf{K} of (5.1.18) as is readily seen from eqs. (5.1.15, 5.1.19). Hence, we have to confine \mathbf{k} to a unit cell in the \mathbf{k} -space spanned by the basic vectors $\mathbf{b}_1, \mathbf{b}_2$ and \mathbf{b}_3 whose volume is given by $\mathbf{b}_1 \cdot (\mathbf{b}_2 \times \mathbf{b}_3) = (2\pi)^3/v_0$. Second, we consider a large crystal lattice spanned by three edges $N\mathbf{a}_1, N\mathbf{a}_2$ and

$N\mathbf{a}_3(N \gg 1)$ and impose cyclic boundary conditions:

$$\xi_{N+1,1,1} = \xi_{1,N+1,1} = \xi_{1,1,N+1} = \xi_{1,1,1}. \quad (5.1.20)$$

Equation (5.1.20) is satisfied by discrete values of \mathbf{k} given by

$$\mathbf{k}_{\ell_1, \ell_2, \ell_3} = (\ell_1/N)\mathbf{b}_1 + (\ell_2/N)\mathbf{b}_2 + (\ell_3/N)\mathbf{b}_3 \quad (5.1.21)$$

with ℓ_i running on the integer values: $0, 1, \dots, N-1$ as readily seen from (5.1.15, 5.1.19).

The choice of the basic \mathbf{k} -region with a more symmetric shape, corresponding to the preference of interval $[-b/2, +b/2]$ to $[0, +b]$ in the one-dimensional lattice, is performed as follows. Consider the planes bisecting those lines connecting the origin $\mathbf{k} = 0$ to all nonvanishing reciprocal lattice points \mathbf{K} given by (5.1.12). The region inside all such planes, the *first Brillouin zone* as it is called, is exactly the basic region to be sought for. Its shape has the highest symmetry so far as it is compatible with the symmetry of the crystal lattice. As for the unit cells and the first Brillouin zones of various types of crystal lattices, see a standard text book on solid state physics.¹ Irrespective of this choice and the type of crystal, the number dn of allowed values of \mathbf{k} within the three-dimensional volume element $d\mathbf{k}$ is given by

$$dn = (V/8\pi^3)dk \quad (5.1.22)$$

where $V = N_1 N_2 N_3 v_0$ is the total volume of the crystal, the same as eq. (1.3.7) for the free space.

5.2 Lattice vibrations in general crystals

If the many identical atoms with interatomic force $v(r)$ mentioned in the preceding section are put in three-dimensional space, they will be arrayed on the lattice which has the greatest number of nearest neighbors, namely either the face-centered-cubic lattice or the hexagonal-close-packed lattice which both have 12 nearest neighbors and, hence, have a binding energy per atom of $6|v(a)|$. Consideration of the second-nearest neighbor interactions or the directional interactions (covalent bonds) would result in a subtle problem concerning the preferment of other lattice structures. Reducing the total binding of the atoms to the sum of pair interactions as described above may be a poor approximation, as is the case for metals where valence electrons are shared by all atoms. The most basic question of solid state physics, why the assemblage of identical atoms chooses a periodic lattice structure as its lowest-energy state, is not yet well-answered theoretically. The situation is hopelessly complicated when there are more than one species of atom.

Leaving elsewhere³ this profound problem which is beyond the scope of this book, we start with the three-dimensional crystal structures as given and generalize

the previous description of vibrational motions of atoms so as to include the case of a unit cell containing more than two atoms. The classical equations of motion for small displacements from their equilibrium positions of the ν th atoms ($\nu = 1, 2, \dots, \sigma$) in the n th unit cell ($n = 1, 2, \dots, N$) are given by

$$M_\nu d^2 \xi_{n\nu} / dt^2 = - \sum_{m\mu} \bar{\bar{f}}_{n\nu, m\mu} \xi_{m\mu} \quad (5.2.1)$$

as can be derived from the harmonic potential

$$U = (1/2) \sum_{n\nu i, m\mu j} f_{n\nu i, m\mu j} \xi_{n\nu i} \xi_{m\mu j} \quad (f_{n\nu i, m\mu j} = f_{m\mu j, n\nu i}) \quad (5.2.2)$$

where i and j refer to the rectangular components of the vector ξ and the force tensor $\bar{\bar{f}}_{n\nu, m\mu}$. The latter can be written as $\bar{\bar{f}}_{\nu\mu}(\mathbf{R}_n - \mathbf{R}_m)$ since it depends on n and m only through the distance $\mathbf{R}_n - \mathbf{R}_m$ between the two lattice points due to the *translational symmetry* of the lattice. One then expects that (5.2.1) has a running wave type solution as in the case of the simple lattice. Instead of ξ and $\bar{\bar{f}}$, however, let us use

$$\eta_{n\nu} \equiv M_\nu^{1/2} \xi_{n\nu}, \quad g_{n\nu i, m\mu j} \equiv (M_\nu M_\mu)^{-1/2} f_{n\nu i, m\mu j}. \quad (5.2.3)$$

Putting $\eta_{n\nu} = \eta_\nu N^{-1/2} \exp[i\mathbf{k} \cdot \mathbf{R}_n - i\omega t + i\phi]$, equations (5.2.1) for $3\sigma N$ variables reduce to those for 3σ variables (the factor three for the components of the vector η):

$$-\omega^2 \eta_\nu + \sum_\mu \bar{\bar{g}}_{\nu\mu}(\mathbf{k}) \eta_\mu = 0 \quad (5.2.4)$$

where

$$\bar{\bar{g}}_{\nu\mu}(\mathbf{k}) \equiv \sum_m \exp[-i\mathbf{k} \cdot (\mathbf{R}_n - \mathbf{R}_m)] \bar{\bar{g}}_{\nu\mu}(\mathbf{R}_n - \mathbf{R}_m), \quad (5.2.5)$$

$$g_{\nu i \mu j}(\mathbf{k}) = g_{\mu j \nu i}^*(\mathbf{k}). \quad (5.2.6)$$

For each wave vector \mathbf{k} , equations (5.2.4) with Hermite matrix $g(3\sigma \times 3\sigma)$ has 3σ real eigenvalues $\omega_s(\mathbf{k})^2$ (non-negative for the lattice to be stable!) with complex eigenvectors $\eta_\nu = \zeta_\nu^{(s, \mathbf{k})}$ ($s = 1, 2, \dots, 3\sigma$) which satisfy the following relations:

$$\sum_{\nu i} \zeta_{\nu i}^{*(s, \mathbf{k})} \zeta_{\nu i}^{(s', \mathbf{k})} = \delta_{s, s'} \quad (\text{orthonormality}), \quad (5.2.7)$$

$$\sum_s \zeta_{\nu i}^{*(s, \mathbf{k})} \zeta_{\mu j}^{(s, \mathbf{k})} = \delta_{\nu\mu} \delta_{ij} \quad (\text{completeness}), \quad (5.2.8)$$

$$\sum_{\nu i \mu j} \zeta_{\nu i}^{*(s, \mathbf{k})} g_{\nu i \mu j}(\mathbf{k}) \zeta_{\mu j}^{(s', \mathbf{k})} = \delta_{s, s'} \omega_s(\mathbf{k})^2 \quad (\text{diagonality}). \quad (5.2.9)$$

The modes are so chosen that angular frequencies $\omega^{(s)}(\mathbf{k})$ (to be taken as ≥ 0) are continous functions of \mathbf{k} for each branch s . The wave vector \mathbf{k} can take N (N is the number of unit cells contained in the macroscopic crystal considered) discrete

values uniformly distributed within the so-called first Brillouin zone of the \mathbf{k} -space. We have thus obtained the normal modes with total number $3\sigma N$ being equal to three times the total number of atoms.

The general solution $\eta_n(t)$ is a linear combination of the normal modes given above:

$$\begin{aligned}\eta_{ni}(t) &= (2N)^{-1/2} \sum_{s,k} A_{s,k} \zeta_{vi}^{(s,k)} \exp(i\mathbf{k} \cdot \mathbf{R}_n - i\omega_{s,k}t) + (c.c.) \\ &= (2N)^{-1/2} \sum_{s,k} [\eta_{s,k}(t) \zeta_{vi}^{(s,k)} + \eta_{(s,-k)}^*(t) \zeta_{vi}^{(s,-k)*}] \exp(i\mathbf{k} \cdot \mathbf{R}_n).\end{aligned}\quad (5.2.10)$$

Here we have introduced the dynamical variables: $\eta_{s,k}(t) = A_{s,k} \exp(-i\omega_{s,k}t)$. Making use of (5.2.7) and the orthogonality

$$N^{-1} \sum_n \exp[i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{R}_n] = \delta_{\mathbf{k},\mathbf{k}'}, \quad (5.2.11)$$

one can calculate the kinetic energy as

$$\begin{aligned}K &= (1/2) \sum_{nvi} (d\eta_{nvi}/dt)^2 \\ &= \sum_{s,k} (\omega_{s,k}^2/4) [2\eta_{s,k}^* \eta_{s,k} - \eta_{s,k} \eta_{s,-k} - \eta_{s,k}^* \eta_{s,-k}^*].\end{aligned}$$

Adding the potential energy U which can be calculated in a similar way with the use of (5.2.9), one finally obtains the Hamiltonian for the harmonic oscillators

$$H = K + U = \sum_{s,k} \omega_{s,k}^2 \eta_{s,k}^* \eta_{s,k} = \sum_{s,k} [P_{s,k}^2 + \omega_{s,k}^2 Q_{s,k}^2]/2 \quad (5.2.12)$$

with the use of *real* coordinates $Q_{s,k}$ and *real* momenta $P_{s,k}$ defined by

$$\eta_{s,k}(t) = [Q_{s,k}(t) + (i/\omega_{s,k})P_{s,k}(t)]/\sqrt{2}. \quad (5.2.13)$$

It is of interest to study the nature of each running wave in eq. (5.2.10) which can be written, in vector form, as

$$A_{s,k} \zeta_v^{(s,k)} \exp(i\mathbf{k} \cdot \mathbf{R}_n - i\omega_{s,k}t). \quad (5.2.14)$$

One can include the phase factor of the complex amplitude $A_{s,k}$ in that of $\zeta_v^{(s,k)}$ since eqs. (5.2.7) to (5.2.9) are invariant against multiplication of $\exp[i\phi^{(s,k)}]$ to $\zeta_v^{(s,k)}$. It is to be noted, however, that the complex vector

$$\zeta_v^{(s,k)} = \zeta_v^{(s,k)} + i\zeta_v^{(s,k)} \quad (5.2.15)$$

cannot be brought into a real vector by multiplying by the scalar phase factor $\exp[i\phi]$ so long as $\zeta_v^{(s,k)}$ and $\zeta_v^{(s,k)}$ are not parallel. Thus the mode (s, \mathbf{k}) in (5.2.10) behaves as a running wave

$$\zeta_v^{(s,k)} \cos(\mathbf{k} \cdot \mathbf{R}_n - \omega_{s,k}t) - \zeta_v^{(s,k)} \sin(\mathbf{k} \cdot \mathbf{R}_n - \omega_{s,k}t). \quad (5.2.16)$$

As time goes on, the atom $n\mathbf{v}$ moves elliptically away from the direction of the axis $\zeta_v^{(s,k)}$ towards $\zeta_v^{(s,k)}$ (Fig. 5.1), and this motion propagates in the direction

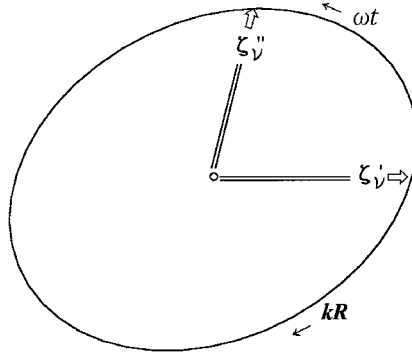


Fig. 5.1 Elliptical rotational motion of normal modes of vibration in crystal lattices with low symmetry. The ellipse shrinks to a line in crystals with higher symmetry.

of \mathbf{k} with the phase velocity $\omega_{s,\mathbf{k}}/k$. For the opposite direction $-\mathbf{k}$, the tensor $g_{\nu\mu}(-\mathbf{k})$ as well as the corresponding eigenvector $\zeta_{\nu}^{*(s,-\mathbf{k})}$ are complex conjugate to those for \mathbf{k} , so that one has to change the sign of the imaginary part $\zeta_{\nu}^{''(s,\mathbf{k})}$. Therefore, the elliptical motion is in the opposite direction for the opposite direction of propagation. Combining this with (5.2.16) and dividing by 2, one obtains the standing wave solution

$$\left[\zeta_{\nu}^{*(s,\mathbf{k})} \cos(\mathbf{k} \cdot \mathbf{R}_n) - \zeta_{\nu}^{''(s,\mathbf{k})} \sin(\mathbf{k} \cdot \mathbf{R}_n) \right] \cos \omega_{s,\mathbf{k}} t \quad (5.2.17)$$

where the spatial pattern rotates *helically* from $\zeta_{\nu}^{*(s,\mathbf{k})}$ to $-\zeta_{\nu}^{''(s,\mathbf{k})}$ as one proceeds towards the direction of \mathbf{k} (Fig. 5.1), and clearly has no inversion symmetry.

The situation is similar to the *optically active* crystals in which the polarization vector of a plane-polarized electromagnetic wave rotates as the wave proceeds. They have crystal structures *without glide symmetry* (see Ref. [11] of Chapter 4), namely it is impossible to attain their mirror-imaged structures by combining any rotation and translation of the original structures, and so they must be classified into right-handed and left-handed helical structures which cannot be brought into registry with each other by any rotation. In fact it can be shown that the tensor $\bar{g}_{\nu\mu}(\mathbf{k})$ defined by (5.2.5), which is invariant against translation of the coordinate \mathbf{R} , is real in crystals with glide symmetry (ZnS structure has this symmetry although it has no inversion symmetry). Namely, the eigenvector $\zeta_{\nu}^{*(s,\mathbf{k})}$ can be taken to be real, and the rotation reduces to a linear oscillation. In crystals without glide symmetry, the imaginary part $\zeta_{\nu}^{''(s,\mathbf{k})}$ increases from zero as k increases from zero. The pitch $\lambda = 2\pi/k$ of the helical pattern decreases from infinity, and tends to a , the pitch of the helical structure of the crystal itself, as k tends to the reciprocal lattice b . Thus the helical pattern of lattice vibrations is *inherently* related to that of the crystal lattice, and is no surprise, although this does not seem to be well known.

However, the situation is somewhat different in the case of optical activity, where the electromagnetic wave is only *weakly affected* by the crystal lattice it traverses due to the *small* radiation–matter interaction, resulting in a *small* rotation of the polarization vector with a pitch *much larger* than its wavelength, as will be shown in Section 6.7.

Let us now study the normal modes of lattice vibrations more elaborately in the limit of long wavelength: $\mathbf{k} \rightarrow 0$, the region of spectroscopic relevance. Electromagnetic waves with angular frequencies resonant to those of typical lattice vibrations: $\omega \sim 10^{13-14} \text{ s}^{-1}$ have wave vectors $k \sim \omega/c \sim 10^5 \text{ m}^{-1}$, which are much smaller than the basic reciprocal lattice, $b \sim 10^{10} \text{ m}^{-1}$, the average radius of the Brillouin zone. Let us note the obvious fact that the translation of all atoms by a *common* value ξ in eq. (5.2.1) causes no force on any atom, indicating that

$$\sum_{m\mu} \bar{\bar{f}}_{nv,m\mu} = 0. \quad (5.2.18)$$

This means that eq. (5.2.4) has a triply-degenerate eigenvalue $\omega^{(s)}(\mathbf{k} \rightarrow 0) = 0$ with three ν -independent eigenvectors $\xi_{vi}^{(s,0)} = \xi_i^{(s,0)}$ (s is assigned the values 1, 2 and 3). They correspond to the *acoustic modes* of lattice vibrations (see the preceding section) in their long-wavelength limit in which all atoms in the same unit cell make a common displacement $\xi(t)$ at any instant. By expanding $\bar{\bar{f}}_{\nu\mu}(\mathbf{k})$ up to the quadratic terms in k_i one can determine the three directions of displacement $\xi^{(s)}$ with the corresponding velocities of propagation defined by $c^{(s)} = \omega^{(s)}(\mathbf{k})/k$ unambiguously as functions of direction \mathbf{k}/k . Velocity $c^{(s)}$ should coincide with the sound-wave velocity calculated with the use of the macroscopic elastic compliance.

The remaining modes ($s = 4, 5, \dots, 3\sigma$) have nonvanishing $\omega^{(s)}(\mathbf{k} = 0)$ in general since the ν -dependent displacements within a unit cell give rise to some restoring forces. They are called the *optical modes*, which may be somewhat misleading since *not all* of them respond to electromagnetic waves (see the next section). For convenience, however, we will follow this convention.

5.3 Interaction of optical modes with electromagnetic field

Different (species of) atoms in a unit cell are differently charged due to their different electron affinities, and hence their relative motions are in general accompanied by electric polarization which responds to electromagnetic waves. We study here how the lattice vibrations respond to the electromagnetic waves. The purpose is twofold: firstly, as a model system for studying the dispersion (frequency dependence) of the dielectric constant, clarifying thereby the meaning of an electric field inside a medium^{4,5} which is essential for a real understanding of the Maxwell equations, and, secondly, as a starting model system for the linear response theory presented

in the next chapter which will provide a better basis of studying other kinds of elementary excitations.

Let us assume the ν th atom of a unit cell is an ion with effective charge e_ν which is an *empirical* parameter to be determined later. Although the total charge of the unit cell vanishes, the displacement ξ_ν of the ion gives rise to dipole moment $e_\nu \xi_\nu$ (e_ν is to be considered as a tensor unless the atom is in an environment of cubic symmetry, as will be mentioned later), and if ξ_ν depends on ν as in the optical modes, the resultant dipolar field, being long ranged, acts upon other atoms in distant unit cells. At a point \mathbf{r} , such dipolar fields from all distant ions, in addition to the external field, are to be included in the macroscopic electric field $\mathbf{E}(\mathbf{r}, t)$ which appears in the Maxwell equations presented in Chapter 1. It may be more practical to consider that the atoms are subject to $\overline{f}_{\nu\mu}^l$, the *short-range* part of the interatomic force, $\overline{f}_{\nu\mu}$, and a long-range dipolar field to be included in $\mathbf{E}(\mathbf{r}, t)$ together with the external field, if any. This *replacement* of \overline{f} by \overline{f}^l has the merit of avoiding possible ambiguity in the convergence in summing the long-range dipolar fields. The related singularity of $\overline{f}(\mathbf{k})$ at $\mathbf{k} = 0$, that it depends upon the direction \mathbf{k}/k even in the long-wavelength limit $k \rightarrow 0$, is thus removed in $\overline{f}^l(\mathbf{k})$. Note also that the replacement does not affect the previous argument on the acoustic modes in their long-wavelength limit because they do not cause a dipolar field. One can now rewrite (5.2.4) as

$$-\omega^2 \eta_\nu + \sum_\mu \overline{g}_{\nu\mu}^l(\mathbf{k}) \eta_\mu = M_\nu^{-1/2} e_\nu \mathbf{E} \quad (\nu = 1, 2, \dots, \sigma) \quad (5.3.1)$$

where \mathbf{E} is to be determined self-consistently with η_ν . The tensor $\overline{g}_{\nu\mu}^l$ is defined in terms of $\overline{f}_{\nu\mu}^l$ by (5.2.3). An important dependence of the lattice vibrations on the direction \mathbf{k}/k in the long-wavelength limit comes from that of \mathbf{E} thus determined, as will be seen. We will solve (5.3.1) in the vanishing limit of \mathbf{k} in the following three steps.⁶ Only in the third step will we consider self-consistency explicitly.

(I) Let us first solve the *homogenous* equation obtained by putting $\mathbf{E} = 0$ in (5.3.1). Its eigenvalue problem was solved formally in Section 5.2. Corresponding to the replacement of \overline{g} by \overline{g}^l , we rewrite the eigenvalues of ω^2 as Ω_s^2 and eigenvectors as $\zeta_v^{(s)}$ ($s = 1, 2, \dots, 3\sigma$) satisfying relations of the same form as (5.2.7) to (5.2.9).

(II) For a *given* electric field $\mathbf{E}(t) = \mathbf{E} \exp(-i\omega t)$, we now study the *forced oscillation* type solution of (5.3.1), $\eta_{\nu\nu}(t) = \eta_\nu(\omega) \exp(-i\omega t)$. One can expand the amplitude $\eta_\nu(\omega)$, a 3σ -dimensional vector, in terms of the *complete* set of 3σ eigenvectors $\zeta_v^{(s)}$, as

$$\eta_{\nu i}(\omega) = \sum_{s'} a^{(s')}(\omega) \zeta_{\nu i}^{(s')}. \quad (5.3.2)$$

By multiplying both sides by $\zeta_{vi}^{(s)}$, summing over (ν, i) and making use of (5.2.7) and (5.2.9), one obtains the equation for determining the expansion coefficients:

$$(\Omega_s^2 - \omega^2)a^{(s)}(\omega) = \mathbf{p}^{(s)} \cdot \mathbf{E}, \quad (5.3.3)$$

$$\mathbf{p}^{(s)} \equiv \sum_{\nu} e_{\nu} \boldsymbol{\xi}_{\nu}^{(s)} \quad (5.3.4)$$

where we have used $\boldsymbol{\eta}_{\nu} = M_{\nu}^{1/2} \boldsymbol{\xi}_{\nu}$. Note that $\mathbf{p}^{(s)} = 0$ for the acoustic mode with ν -independent $\boldsymbol{\xi}_{\nu}^{(s)}$ due to electric neutrality. With the use of (5.2.8) and (5.3.4), one can derive the following sum rule:

$$\sum_s p_i^{(s)} p_j^{(s)} = \left[\sum_{\nu} e_{\nu}^2 / M_{\nu} \right]_{ij}. \quad (5.3.5)$$

Let us note that the macroscopic polarization \mathbf{P} consists of an *electronic* contribution – the field-induced changes in the electronic wave functions of the ions – and an *ionic* contribution – the displacements of the ions:

$$\mathbf{P} = (\epsilon_e - \epsilon_0)\mathbf{E} + N_0 \sum_{\nu} e_{\nu} \boldsymbol{\xi}_{\nu} \quad (5.3.6)$$

where N_0 is the number of unit cells in unit volume. As shown schematically in Fig. 5.2 for a dielectric dispersion in a typical insulator with bandgap energy much greater than the phonon energies (a condition which does not apply in narrow-gap semiconductors and metals), ϵ_e is the notation for the dielectric constant $\epsilon(\omega)$ in that region of frequency ω which is (1) *much higher* than the lattice vibrational frequencies but (2) *well below* the electronic bandgap ϵ_g divided by \hbar . (Because of (1) ϵ_e has often been called the high-frequency dielectric constant and written

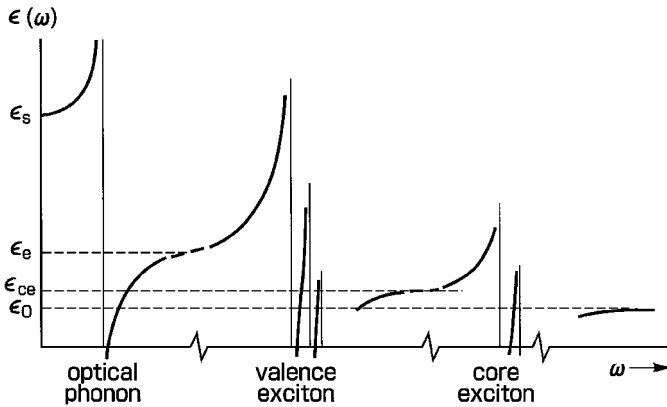


Fig. 5.2 Typical dispersion of dielectric function $\epsilon(\omega)$ of an insulator. At frequencies far higher than those of optical phonons but a little below the bandgap appear dispersions due to excitons associated with a valence-band hole (see Chapter 8). At even higher frequencies appear dispersions due to excitons associated with a core hole (see Chapter 12).

as ϵ_∞ which, however, is confusing because of (2); in fact, $\epsilon(\omega \rightarrow \infty) = \epsilon_0$, the vacuum dielectric constant.) In this *optically* transparent ω -region of an insulator, which usually overlaps with the visible region (ϵ_e has often been called the *optical* dielectric constant for this reason, but its abbreviated notation ϵ_o , which has also been used, is in danger of being confused both with ϵ_0 and $\epsilon(\omega \rightarrow 0) \equiv \epsilon_s$, the *static* dielectric constant), only the *electronic* polarization (due to electrons bound within atoms, filling up the valence band just below the energy gap as well as the inner atomic core states) can follow the oscillating electric field instantaneously (whence the suffix of ϵ_e used in this book).

Since \mathbf{E} and thereby the forced $\boldsymbol{\xi}$ s oscillate as $\exp(-i\omega t)$ in (5.3.6), the same is true for \mathbf{P} and the electric flux density $\mathbf{D} = \epsilon \mathbf{E} = \epsilon_0 \mathbf{E} + \mathbf{P}$. With the use of this equation and eqs. (5.3.3) to (5.3.5), the dielectric constant ϵ as a function of ω – the *dielectric function* – is given by

$$\epsilon(\omega) = \epsilon_e + N_0 \sum_s \mathbf{p}^{(s)} \mathbf{p}^{(s)} / (\Omega_s^2 - \omega^2), \quad (5.3.7)$$

where $\mathbf{p}^{(s)} \mathbf{p}^{(s)}$ represents a symmetric tensor whose (i j)-component is given by $p_i^{(s)} p_j^{(s)}$. Likewise, the dielectric function $\epsilon(\omega)$, its high-frequency value ϵ_e as well as the effective charge e_v which are proportionality constants relating two vector quantities, are symmetric *tensors*, which reduce to *scalars* only in those crystals (for macroscopic quantities such as ϵ) or environments (for microscopic quantities such as e_v) with *cubic symmetry*.

(III) We now consider self-consistent solutions for the lattice vibrations. While the general relations derived in (II) hold irrespective of the origin of \mathbf{E} , we now specify \mathbf{E} to be purely of internal origin; namely, we consider only the depolarizing field given by

$$0 = \rho = \text{div} \mathbf{D} = i\mathbf{k} \cdot (\epsilon_0 \mathbf{E} + \mathbf{P}) \quad (5.3.8)$$

for the polarization wave of the form $\exp(i\mathbf{k} \cdot \mathbf{r} - i\omega t)$ ($\mathbf{k} \rightarrow 0$). Since $\nabla \times \mathbf{E} = 0$ under the neglect of the retardation effect (take the limit of $c \rightarrow \infty$ in eqs. (1.1.1, 1.1.2)), \mathbf{E} is a longitudinal field given by

$$\epsilon_0 \mathbf{E} = -\mathbf{k}(\mathbf{k} \cdot \mathbf{P}), \quad \mathbf{k} \equiv \mathbf{k}/k. \quad (5.3.9)$$

Combining this with the relation $\mathbf{P} = (\epsilon(\omega) - \epsilon_0)\mathbf{E}$, one obtains

$$[\epsilon_0 + \{(\epsilon(\omega) - \epsilon_0)\mathbf{k}\} \mathbf{k}] \mathbf{P} = 0 \quad (5.3.10)$$

which assures that the lattice vibration has been solved *self-consistently* under the depolarizing field caused by itself. In a coordinate system with z -axis taken parallel to \mathbf{k} , eq. (5.3.10) can be written as

$$\epsilon_0 P_{x'} + \epsilon_{x'z'} P_{z'} = 0, \quad \epsilon_0 P_{y'} + \epsilon_{y'z'} P_{z'} = 0, \quad \epsilon_{z'z'} P_{z'} = 0, \quad (5.3.10a)$$

A non-trivial ($\mathbf{P} \neq 0$) solution of (5.3.10) exists only when $\epsilon_{\parallel} \equiv \epsilon_{z'z'}$, the determinant for the tensor in $[\cdots]$ divided by ϵ_0^2 , vanishes, namely, at ω satisfying

$$\epsilon_{\parallel}(\omega) = \epsilon_{e\parallel} + N_0 \sum_s (p_{\parallel}^{(s)})^2 / (\Omega_s^2 - \omega^2) = 0. \quad (5.3.11)$$

Among the optical modes ($s = 4, 5, \dots, 3\sigma$), those with nonvanishing $\mathbf{p}^{(s)}$ are called *infrared active* since they interact with and absorb the electromagnetic waves, mostly in the infrared region. The remaining modes, with vanishing $\mathbf{p}^{(s)}$ because of the symmetry of the crystal structure, are *infrared inactive*. Here the Ω_s 's themselves give the eigenfrequencies of the lattice vibrations irrespective of direction \mathbf{k} since they have been obtained by ignoring the depolarizing field. (Note that a zero point of $\epsilon_{\parallel}(\omega)$ tends to the pole Ω_s in the vanishing limit of $\mathbf{p}^{(s)}$.) Most of these infrared-inactive modes can be observed by Raman scattering (*Raman active modes*) in which incident light with ω is scattered to $\omega \mp \Omega_s$ (*Stokes* and *anti-Stokes* lines) through the dependence of the polarizability on the atomic displacements.⁷

Let us study the infrared-active modes in more detail. Renumbering them as $a = 1, 2, \dots, h$ ($\leq 3\sigma - 3$), we show the dielectric dispersion (5.3.11) schematically in Fig. 5.3(a), in which \mathbf{k} -independent poles Ω_a and \mathbf{k} -dependent zeros $\omega_a(\mathbf{k})$ appear

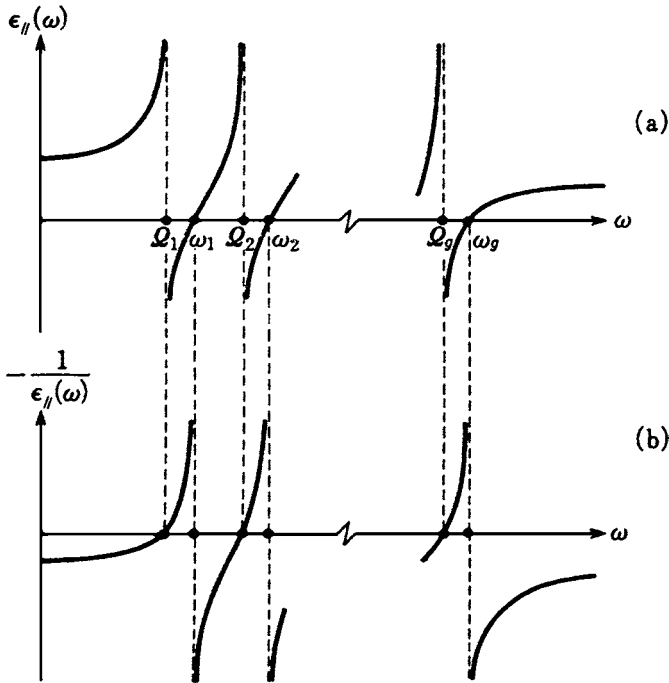


Fig. 5.3 (a) Dispersion of $\epsilon_{\parallel}(\omega)$ in anisotropic crystal in which poles Ω_s and zeros ω_s (eigenfrequencies) appear alternately. (b) Dispersion of $-1/\epsilon_{\parallel}(\omega)$ with zeros Ω_s and poles ω_s .

alternately. Noting that $\epsilon(\omega)$ tends to ϵ_e after ω has exceeded the highest Ω_h , one can factorize (5.3.11) as

$$\epsilon_{\parallel}(\omega)/\epsilon_e = \prod_a [\omega_a(\mathbf{k})^2 - \omega^2] / (\Omega_a^2 - \omega^2). \quad (5.3.12)$$

Take the x -axis along $\mathbf{p}^{(a)}$, the z -axis and x' -axis within the $z'(\parallel \mathbf{k})$ - x plane keeping the y -axis unchanged, and denote the angle between z' and x by θ . As \mathbf{k} tends perpendicular to $\mathbf{p}^{(a)}$ (namely, $\cos \theta \rightarrow 0$), $\omega_a(\mathbf{k})$ (or $\omega_{a-1}(\mathbf{k})$) $\rightarrow \Omega_a$ tends to 0 as $\cos^2 \theta$ while the dominant term of $\epsilon_{x'z'} \propto p_x^{(a)} p_{z'}^{(a)} / (\Omega_a^2 - \omega_a^2) \propto (\cos \theta)^{-1}$ and hence $P_{z'}/P_{x'} \propto \cos \theta$, namely, \mathbf{P} and \mathbf{k} tend parallel to and perpendicular to $\mathbf{p}^{(a)}$, respectively. This means that all modes of lattice vibrations with frequency Ω_a and with \mathbf{k} perpendicular to $\mathbf{p}^{(a)}$ are transverse waves of macroscopic polarization \mathbf{P} (but not necessarily of the displacements of individual ions). This is quite natural since the difference $\omega_a(\mathbf{k})^2 - \Omega_a^2$ originates in the depolarizing field.

A study of a few examples of differing crystal structures will be of practical use. In uniaxial crystals such as CdS, some Ω_a s are doubly degenerate with two independent $\mathbf{p}^{(a)}$ s of equal length which are perpendicular to the principal c -axis and to each other, while others are non-degenerate with $\mathbf{p}^{(a)}$ parallel to the c -axis. The former case can be considered as a limiting case of $\Omega_{a+1} \rightarrow \Omega_a$. The modes of lattice vibration with ω_a sandwiched between them tend to transverse waves for any \mathbf{k} (note that one can always find $\mathbf{p}^{(a)}$ perpendicular to both \mathbf{k} and the principal c -axis) in the limit. Other modes with $\omega_{a'}(a' = a + 1 \text{ or } a - 1)$ are mixed modes which tend to transverse waves as \mathbf{k} tends parallel to the c -axis and $\omega_{a'}$ tends to $\Omega_a = \Omega_{a+1}$. The proof is left to the reader.

Another important example is an isotropic crystal with cubic symmetry, of which the Ω_a s are always triply degenerate with three mutually perpendicular vectors $\mathbf{p}^{(a)}$, $\mathbf{p}^{(a+1)}$ and $\mathbf{p}^{(a+2)}$ which can be taken arbitrarily to the crystalline axis. One can readily find that $\omega_a = \omega_{a+1} \equiv \omega_{at}(=\Omega_a)$ represent transverse waves while $\omega_{a+2} \equiv \omega_{al}(> \Omega_a)$ represent longitudinal waves for any \mathbf{k} . The proof is again left to the reader. Renumbering a so as to represent each set of the triply degenerate Ω_a and noting that ϵ is a scalar here, one can rewrite eq. (5.3.12) as

$$\epsilon(\omega)/\epsilon_e = \prod_a (\omega_{at}^2 - \omega^2) / (\omega_{at}^2 - \omega^2). \quad (5.3.13)$$

Isotropic ionic crystals with a pair of ions in a unit cell, such as the NaCl type and the CsCl type, have one set of infrared-active modes. Denoting the “effective” charges of a cation and an anion by $\pm e'$ and their reduced mass by $M_r(M_r^{-1} \equiv M_+^{-1} + M_-^{-1})$, one obtains $|\mathbf{p}^{(a)}| = e'/\sqrt{M_r}$ with the use of the sum rule (5.3.5). Thus one can rewrite (5.3.7) as

$$\epsilon(\omega) = \epsilon_e + (N_0 e'^2 / M_r) / (\omega_t^2 - \omega^2) = \epsilon_e + (\epsilon_s - \epsilon_e) \omega_t^2 / (\omega_t^2 - \omega^2) \quad (5.3.14)$$

where $\epsilon_s \equiv \epsilon(\omega = 0)$ is the *static* dielectric constant. By putting $\omega = 0$ in (5.3.14), one obtains

$$\omega_\ell^2 = \omega_t^2 + N_0 e'^2 / \epsilon_e M_r, \quad (5.3.15)$$

$$\omega_\ell^2 / \omega_t^2 = \epsilon_s / \epsilon_e. \quad (5.3.16)$$

The second term of (5.3.15) originates from the depolarizing field caused by the relative displacement of the ions, the same as that caused by the plasma oscillation of the electron gas in metals and the ionized gases which will be mentioned later. The relation (5.3.16), connecting the longitudinal and transverse frequencies with the dielectric constants at appropriate frequencies was generalized by Cochran and Cowley⁸ and Kurosawa,⁶ as is obtained by putting $\omega = 0$ in (5.3.13). The values of the effective charge “ e' ” introduced above can be obtained from the analysis of the observed dielectric dispersion, and turn out to be a few tens percent away from the integer values simply expected for monovalent, divalent, . . . ions.

Finally, we make some remark on the effective charge e_v , which was introduced at the beginning of this section and *defined* explicitly in (5.3.6) as a tensor relating the polarization to the atomic displacements. Although it was used earlier in (5.3.1), the force on its r.h.s. can be derived from the definition as follows. The work done by the electric field to polarize matter can be written, according to (5.3.6), as $dW = \mathbf{E} \cdot d\mathbf{P} = N_0 \sum_{vij} E_i e_{vij} d\xi_{vj} = N_0 \sum_{vij} d\xi_{vi} e_{vij} E_j$ (we have used the fact that e_v is a symmetric tensor as ϵ is, the properties derivable from energy conservation), whence the force on the ion v is given by $N_0^{-1} \partial W / \partial \xi_{vi} = \sum_j e_{vij} E_j$, as shown in (5.3.1). This definition, though logically natural and practically useful, does not necessarily fit a naive interpretation. As one example, the effective charge e' of ions in alkali halides which are supposed to be ideal monovalent ionic crystals, determined by (5.3.15) with the use of the measured values of ω_t , ω_ℓ and ϵ_∞ , deviates significantly (though not too much) from the electronic charge $|e|$. As a second example, a unit cell of the triclinic Se crystal consists of three Se atoms in *non-equivalent* sites, because of which they have finite effective charges, giving rise to an infrared-active mode. In contrast, a unit cell of diamond consists of two carbon atoms occupying *equivalent* sites, which is the reason for the absence of infrared-active mode.

5.4 Localized and quasi-localized modes

If a host atom of the crystal lattice is replaced by an impurity atom, the lattice vibrations are perturbed in its neighborhood, and there can arise a localized mode with ω as a discrete spectrum outside the continuous frequency spectrum of the unperturbed lattice vibrations. To obtain basic concepts, we consider the simplest situation in which only the change of mass ($M \rightarrow M' = M + \Delta M$) but not of the

force constants of the atom on the site $n = 0$ of a monatomic lattice is taken into account. Assuming that only the impurity atom has charge e_0 , the equations of motion for the atoms under an alternating electric field are written as

$$(M + \delta_{n0}\Delta M)d^2\xi_n/dt^2 = -\sum_m \bar{f}(\mathbf{R}_n - \mathbf{R}_m)\xi_m(t) + \delta_{n0}e_0\mathbf{E}\exp(-i\omega t). \quad (5.4.1)$$

As was shown in Section 5.2, the host lattice has normal modes given by

$$\xi_n^{(s,k)} = N^{-1/2}\xi^{(s,k)}\exp(i\mathbf{k} \cdot \mathbf{R}_n), \quad (5.4.2)$$

$$M\sum_n \xi_n^{*(s,k)} \cdot \xi_n^{(s',k')} = M\xi^{*(s,k)} \cdot \xi^{(s',k')}\delta_{kk'} = \delta_{ss'}\delta_{kk'}, \quad (5.4.3)$$

$$\omega_{s,k}^2 = \xi^{*(s,k)}\bar{f}(\mathbf{k})\xi^{(s,k)}, \quad \bar{f}(\mathbf{k}) \equiv \sum_m \bar{f}(\mathbf{R}_n - \mathbf{R}_m)\exp[-i\mathbf{k} \cdot (\mathbf{R}_n - \mathbf{R}_m)]. \quad (5.4.4)$$

One can expand the amplitude of the forced oscillation solution of (5.4.1),

$$\xi_n(t) = \xi_n \exp(-i\omega t), \quad (5.4.5)$$

in terms of the complete set (5.4.2) as

$$\xi_n = \sum_{s',k'} a_{s',k'} \xi_n^{(s',k')}. \quad (5.4.6)$$

Inserting (5.4.5, 5.4.6) into (5.4.1) and multiplying it by $\sum_n \xi_n^{*(s,k)}$, one obtains the equations for the expansion coefficients:

$$(\omega_{s,k}^2 - \omega^2)a_{s,k} = N^{-1/2}\xi^{*(s,k)} \cdot [\Delta M/M)\omega^2\xi_0 + (e_0/M)\mathbf{E}]. \quad (5.4.7)$$

With the use of (5.4.2, 5.4.6 and 5.4.7), one obtains a self-consistent equation for ξ_0 and \mathbf{E} :

$$\xi_0 = \bar{G}(\omega^2)[(\Delta M/M)\omega^2\xi_0 + (e_0/M)\mathbf{E}] \quad (5.4.8)$$

where $\bar{G}(\omega^2)$ is a tensor which reduces to a scalar in the cubic lattice:

$$\bar{G}(\omega^2) \equiv N^{-1} \sum_{s,k} M\xi^{(s,k)}\xi^{*(s,k)}(\omega_{s,k}^2 - \omega^2)^{-1} \quad (5.4.9)$$

$$= (3N)^{-1} \sum_{s,k} (\omega_{s,k}^2 - \omega^2)^{-1} \quad (\text{cubic}). \quad (5.4.9a)$$

In deriving the latter, one has only to calculate the trace of the tensor with the use of (5.4.3)

$$M \text{ trace}[\xi^{(s,k)}\xi^{*(s,k)}] = M \sum_i \xi_i^{(s,k)}\xi_i^{*(s,k)} = 1$$

and then apply the cubic symmetry.

In order to describe appropriately the dissipation when ω of the incoming electromagnetic wave is inside the continuous spectrum of $\omega_{s,k}$ of the lattice vibrations, let us add an infinitesimal imaginary part $i\gamma$ ($\gamma \rightarrow +0$) to ω , leaving a more sound basis of argument to Chapter 6. Then one can *redefine* the complex function $G(\omega^2)$ analytic in the upper half of the complex plane of ω :

$$\begin{aligned} G(\omega^2) &\equiv \lim_{\gamma \rightarrow 0} G((\omega + i\gamma)^2) \equiv G'(\omega^2) + i(G''(\omega^2)) \\ &= (3N)^{-1} \sum_{s,k} [P(\omega_{s,k}^2 - \omega^2)^{-1} + i\pi \operatorname{sgn}(\omega) \delta(\omega_{s,k}^2 - \omega^2)] \end{aligned} \quad (5.4.10)$$

where P is the principal value and $\operatorname{sgn}(\omega) \equiv \pm 1$ according as $\omega >$ or < 0 .

The energy absorbed per time by the impurity atom, divided by the energy density of the electromagnetic wave, $|E|^2$, namely, the absorption rate, is given, with the use of the solution of (5.4.8) for ξ_0 and $d\xi_0/dt = -i\omega\xi_0$, by

$$\begin{aligned} W(\omega) &= \operatorname{Re} [(d\xi_0/dt)e_0 E^*]/|E|^2 = [e_0^2/M] 2\omega G''(\omega^2) \\ &\times \left[\{1 - (\Delta M/M)\omega^2 G(\omega^2)\}_I^2 + \{(\Delta M/M)\omega^2 G''(\omega^2)\}_I^2 \right]^{-1} \end{aligned} \quad (5.4.11)$$

The calculated spectra of $W(\omega)/\omega$ are shown in Fig. 5.4⁹ for the simple cubic lattice with the nearest neighbor force f and hence with $\omega_{s,k}^2 = (2f/M)[3 - \sum_{xyz} \cos(k_x a)]$. The abscissa is chosen as $\omega^2/(2f/M)$. With $\Delta M = 0$, the normalized density $D(\omega^2) = (3N)^{-1} \sum_{s,k} \delta(\omega_{s,k}^2 - \omega^2) = \pi^{-1} G''(\omega^2)$ of the host lattice vibrations is directly reflected on the differential absorption spectrum

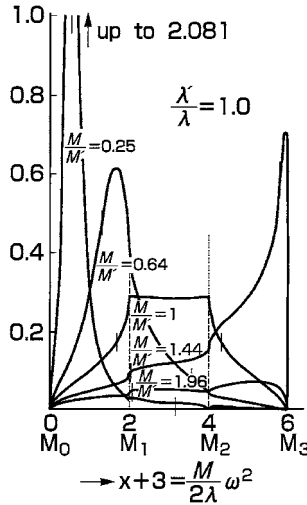


Fig. 5.4 Calculated absorption spectra, $W(\omega)/\omega$, due to a charged impurity atom with mass M' substituted for a host atom with mass M whereby the force constant is assumed not to change. In the case of $M/M' = 1.94$, a line spectrum due to a localized mode appears (not shown in the figure) above the upper end of the continuum of the host lattice vibrations so as to conserve the integrated area of the spectra.⁹

$[W(\omega)/\omega]d(\omega^2) = 2W(\omega)d\omega$ through the charge e_0 of the impurity atom (see the symmetric curve with $M/M' = 1$ in Fig. 5.4). With $\Delta M \neq 0$, the lattice vibrations are distorted near the impurity atom, giving rise to the distortion of the absorption spectrum. When the impurity mass $M' = M + \Delta M$ is smaller than a certain value, which is 0.69 in this case, there appears a *line spectrum* due to a *localized mode* at a frequency higher than the continuous spectrum of the extended modes, at the position $\omega = \omega_{\text{loc}}$ where the expression in $\{\cdots\}_{\text{I}}$ of eq. (5.4.11) vanishes. (In fact, by expanding $\{\cdots\}_{\text{I}} = c(\omega^2 - \omega_{\text{loc}}^2) + \cdots$ one finds that (5.4.11) gives a Lorentzian function of ω^2 with width proportional to $G''(\omega)^2$ which is *infinitesimally small* here outside the continuum.) As $M + \Delta M$ increases, ω_{loc} appears in the continuum, giving rise to a peak with finite width due to the *finite* value of $\{\cdots\}_{\text{II}}$, though the peak is only significant near both ends of the continuum where the density $D(\omega^2)$ is small. The peak can be interpreted as a *quasi-localized mode* with *finite lifetime*. A more exact statement is that the spacially extended lattice vibrations with ω near ω_{loc} have *enhanced* amplitudes at the impurity atom, and hence respond more strongly to the electromagnetic wave. However, the total area of the absorption spectrum, inclusive of the discrete line if any, satisfies the sum rule

$$\int_0^\infty d(\omega^2)W(\omega^2) = \pi e_0^2/(M + \Delta M), \quad (5.4.12)$$

as it should. (The proof of this with the use of contour integration on the complex plane of ω is left to the reader.) The result is consistent with eq. (5.3.5).

5.5 Quantum mechanics of lattice vibrations

It is straightforward to quantize the harmonic lattice vibrations. Similarly to the case of electromagnetic waves, one has only to consider the *real* dynamical variables $Q_{s,\mathbf{k}}$ and $P_{s,\mathbf{k}}$ introduced by (5.2.13) as the quantal coordinate and momentum (although they do not represent the pure coordinate and momentum) of the harmonic oscillator (s, \mathbf{k}) or, if more convenient, to introduce the annihilation and creation operators $b_{s,\mathbf{k}}$ and $b_{s,\mathbf{k}}^\dagger$ defined by (see eq. (2.2.4))

$$b_{s,\mathbf{k}}, b_{s,\mathbf{k}}^\dagger \equiv (\omega_{s,\mathbf{k}}/2\hbar)^{1/2} [Q_{s,\mathbf{k}} \pm i P_{s,\mathbf{k}}/\omega_{s,\mathbf{k}}]. \quad (5.5.1)$$

The displacement of the ν th atom in the n th unit cell is expressed by

$$\begin{aligned} \xi_{n\nu} = & (2NM_\nu)^{-1/2} \sum_{s,\mathbf{k}} (\hbar/\omega_{s,\mathbf{k}})^{1/2} \\ & \times [b_{s,\mathbf{k}} \zeta_\nu^{(s,\mathbf{k})} \exp(i\mathbf{k} \cdot \mathbf{R}_n) + b_{s,\mathbf{k}}^\dagger \zeta_\nu^{*(s,\mathbf{k})} \exp(-i\mathbf{k} \cdot \mathbf{R}_n)], \end{aligned} \quad (5.5.2)$$

whereas the Hamiltonian for the lattice vibrations is given by

$$H = \sum_{s,k} \hbar \omega_{s,k} \left(b_{s,k}^\dagger b_{s,k} + \frac{1}{2} \right). \quad (5.5.3)$$

The energy quantum $\hbar \omega_{s,k}$ of the mode (s, k) of the lattice vibration is called a *phonon* irrespective of whether the mode is acoustic (the origin of the word) or optical.

The quantal nature of lattice vibrations was first noted by Einstein (1907) to explain the anomalous decrease at low temperature of the specific heat of solids. This was soon after his proposal of the quantal nature of light (1905) to explain the photoelectric effect. The success of these studies on two different systems lent strong support to Planck's working hypothesis, making an important contribution to the discovery of quantum mechanics. Similarly to the thermal radiation field, the thermal energy of lattice vibrations is given by

$$E(T) = \sum_{s,k} \langle \langle E_{s,k} \rangle \rangle = \sum_{s,k} \hbar \omega_{s,k} \left[\left(\frac{1}{2} \right) + \{ \exp(\beta \hbar \omega_{s,k}) - 1 \}^{-1} \right]. \quad (5.5.4)$$

In contradistinction to the case of the radiation field, the zero-point vibrational energy (the term $(\frac{1}{2})$ in $[\cdot \cdot \cdot]$) of the lattice vibration is an observable quantity which in fact decreases the binding energy of a solid and even prevents solidification of light atoms such as helium under no applied pressure. At high temperatures each mode contributes energy $k_B T$, in agreement with the result of classical statistical mechanics and hence with the empirical rule due to Dulong and Petit that the specific heat of a solid is $3k_B$ per atom. The low-temperature behaviors of the contribution to the specific heat are quite different between the acoustic modes with vanishing energies with $k \rightarrow 0$ and the optical modes with nonvanishing energies. The specific heat rises as T^3 in the former modes but much more slowly in the latter, as is well known.

The anharmonic part of the interatomic potentials (the third and higher orders in b and/or b^\dagger), which we have ignored, causes the interaction among phonons which is absent in the harmonic approximation, giving rise to a variety of scattering processes in which three or more phonons participate either by absorption (b) or emission (b^\dagger). In order that the scattering process is completed as a real process, the energy should be conserved between the initial and final states, but the wave vector k ($\hbar k$ is to be distinguished from real momentum) needs only to satisfy the conservation rule with the possible participation of the reciprocal lattice \mathbf{K} (see eq. (5.1.18)) in order to be compatible with the translational symmetry of

the lattice. For instance, in three-phonon processes in which a phonon \mathbf{k} and a phonon \mathbf{k}' collide and are converted to a phonon \mathbf{k}'' , the wave vector conservation is given by

$$\mathbf{k} + \mathbf{k}' = \mathbf{k}'' + \mathbf{K}. \quad (5.5.5)$$

The process with nonvanishing \mathbf{K} is called the *Umklapp process*, and plays an important role in keeping thermal conduction by phonons to a finite value.

6

Electric susceptibility and dielectric constant

6.1 Linear response theory for electric susceptibility

In the preceding chapter, we introduced the dielectric constant for *harmonic* lattice vibrations for which the response (polarization) was necessarily *linear* in an external electromagnetic field. That is not the case in general systems such as anharmonic vibrational systems and electronic systems. Nevertheless, for a *small* external field, one expects that the responses of the system are *linear* in the field, and in fact one can give general expressions for the proportionality coefficients, such as susceptibilities and transport coefficients, of a general system which otherwise is in thermal equilibrium, provided the Hamiltonians of the system and its interaction with the external field are explicitly given.^{1,2}

For example, application of a small electric field \mathbf{E} gives rise to small electric polarization \mathbf{P} and a small electric current $\mathbf{J} = d\mathbf{P}/dt$ in the system, the coefficients being electric susceptibility χ and electric conductivity σ , respectively. When the external field varies with position and time, so does the response, which one can write as

$$\mathbf{P}(\mathbf{r}, t) = \int d\mathbf{r}' \int dt' \chi(\mathbf{r}, t; \mathbf{r}', t') \mathbf{E}(\mathbf{r}', t') \quad (6.1.1)$$

with the use of the spacetime-dependent response function for the susceptibility tensor χ , since the linearity assures the superposition principle: the sum of the causes leads to the sum of the corresponding results. Since the result can never precede the cause, one has the relation:

$$\chi(\mathbf{r}, t; \mathbf{r}', t') = 0 \quad (t < t') \quad (\text{causality}). \quad (6.1.2)$$

(Note that we have ignored the relativistic requirement that χ vanishes for $|\mathbf{r} - \mathbf{r}'| > c_0(t - t')$, assuming the light velocity to be infinite.)

If one is concerned with macroscopic quantities in macroscopic systems in thermal equilibrium, neglecting microscopic fluctuations over microscopic times, one

also expects translational symmetry in time:

$$\chi(\mathbf{r}, t; \mathbf{r}', t') = \chi(\mathbf{r}; \mathbf{r}', t - t') \quad (\text{stationarity}). \quad (6.1.3)$$

If the system can be considered spacially homogeneous, one also expects translational symmetry in space:

$$\chi(\mathbf{r}, t; \mathbf{r}', t') = \chi(\mathbf{r} - \mathbf{r}', t - t') \quad (\text{homogeneity}). \quad (6.1.4)$$

Since crystals have translational symmetry with finite lattice constants as periods, (6.1.4) is meaningful only when one is concerned with the position \mathbf{r} on the *macroscopic* scale.

Let us now derive the microscopic expression for the macroscopic linear response coefficient χ , for given Hamiltonians H for the unperturbed system and H_t' for its interaction with a time-dependent external field. The interaction of matter with an electromagnetic field was given by eqs. (3.3.7, 3.3.10), which we write

$$H_t' = - \int d\mathbf{r} \mathbf{j}(\mathbf{r}) \cdot \mathbf{A}(\mathbf{r}, t), \quad (6.1.5)$$

$$\mathbf{j}(\mathbf{r}) \equiv \sum_i (e_i/m_i)[\mathbf{p}_i \delta(\mathbf{r}_i - \mathbf{r}) + \delta(\mathbf{r}_i - \mathbf{r}) \mathbf{p}_i]/2 \quad (6.1.6)$$

for our program in which we take the vector potential $\mathbf{A}(\mathbf{r}, t)$ as a given external field which should be treated as a classical quantity (c-number), not the quantum field (q-number) as was given in (3.3.7). However, the current density $\mathbf{j}(\mathbf{r})$ of the “system” should be treated quantum-mechanically, and hence should be symmetrized with respect to non-commuting position \mathbf{r}_i and momentum \mathbf{p}_i so as to be an Hermite operator, while \mathbf{r} is simply the macroscopic position (c-number) as a function of which we define the fields \mathbf{A} and \mathbf{j} .

Let us solve the time-dependent Schrödinger equation for the perturbed system,

$$i\hbar \partial \Psi / \partial t = (H + H_t') \Psi \quad (6.1.7)$$

with the help of the *interaction representation* in which the unperturbed Hamiltonian H is removed as follows. By putting

$$\Psi(t) = \exp(-itH/\hbar) \Psi'(t) \quad (6.1.8)$$

and inserting it into (6.1.7), one obtains the Schrödinger equation for $\Psi'(t)$:

$$i\hbar \partial \Psi' / \partial t = H_t'(t) \Psi' \quad (6.1.9)$$

with the use of the Heisenberg representation (see eq. (2.1.18)):

$$H_t'(t) \equiv \exp(itH/\hbar) H_t' \exp(-itH/\hbar) \quad (6.1.10)$$

of H_t' defined in terms of the unperturbed H . The interaction representation, (6.1.8), is so to speak the mixture of the Schrödinger and Heisenberg representations.

Denote the eigenvalues and eigenstates of the system Hamiltonian H by E_ℓ and ϕ_ℓ and expand $\Psi(t)$ in terms of the set $\{\phi_\ell\}$ as

$$\Psi(t) = \sum_\ell c_\ell(t) \phi_\ell. \quad (6.1.11)$$

Putting (6.1.11) into (6.1.9), multiplying ϕ_ℓ^* and integrating, one obtains

$$i\hbar dc_\ell(t)/dt = \sum_{\ell'} (H'_\ell(t))_{\ell\ell'} c_{\ell'}(t). \quad (6.1.12)$$

Starting from the initial state ℓ at $t = -\infty$, the wave function of the system at time t is given, up to the first order in H'_ℓ , by

$$\Psi'_\ell(t) = \phi_\ell + \sum_{\ell'}' \phi_{\ell'} (i\hbar)^{-1} \int_{-\infty}^t dt' (H'_\ell(t'))_{\ell\ell'} \quad (6.1.13)$$

where the prime' of \sum' indicates that the term with $\ell' = \ell$ is to be omitted.

Let us now define the operators for electric polarization and current density as functions of the *macroscopic* position coordinate \mathbf{r} (c-number) by

$$\mathbf{P}(\mathbf{r}) \equiv \sum_i e_i \mathbf{r}_i \Lambda(\mathbf{r} - \mathbf{r}_i), \quad (6.1.14)$$

$$\mathbf{J}(\mathbf{r}) \equiv \sum_i e_i [\mathbf{p}_i \Lambda(\mathbf{r} - \mathbf{r}_i) + \Lambda(\mathbf{r} - \mathbf{r}_i) \mathbf{p}_i] / 2, \quad (6.1.15)$$

where the summation is to be extended on all charged particles with charge e_i and position coordinate \mathbf{r}_i (operator) of the material system, and the *locator* $\Lambda(\mathbf{r})$ is a *positive smooth* function which is nonvanishing only within a *macroscopically small* region: $r < a_M$ (defined in (6.1.17) below) and satisfies

$$\int d\mathbf{r} \Lambda(\mathbf{r}) = 1. \quad (6.1.16)$$

The symmetrization in (6.1.15) is needed for $\mathbf{J}(\mathbf{r})$ to be an Hermite operator. It is an elaboration of the earlier definition (3.3.10). Equation (6.1.14) is a natural extension of the dipole moment $\boldsymbol{\mu}$, defined by (3.4.10), of a molecular system of microscopic size a_m to that of a material system of macroscopic size in the form of *density* – dipole moment per unit volume at a macroscopically specified position \mathbf{r} . The spatial integration of (6.1.14 and 6.1.15) correctly gives the total moment and the total current, respectively, due to the normalization (6.1.16). On the other hand, we are going to describe the response of matter to an external field which spatially varies smoothly in the macroscopic region, as is more explicitly specified by the wavelength λ in the case of an electromagnetic wave. Namely, we have to define the position coordinate \mathbf{r} with an accuracy: $|\delta\mathbf{r}| \ll \lambda$ in order that our space-dependent description of $\mathbf{P}(\mathbf{r})$ against the external field $\mathbf{A}(\mathbf{r}')$ is meaningful. Thus we have to impose the following restriction on what we call the macroscopically small size a_M :

$$a_m \ll a_M \ll \lambda. \quad (6.1.17)$$

The expectation value of the polarization density in the state (6.1.13) is then given by

$$\begin{aligned}
 (\Psi_\ell(t), \mathbf{P}(\mathbf{r})\Psi_\ell(t)) &= (\Psi_\ell^l(t), \mathbf{P}(\mathbf{r}, t)\Psi_\ell^l(t)) = \mathbf{P}(\mathbf{r})_{\ell\ell} \\
 &+ \sum (i\hbar)^{-1} \sum_{\ell'} (\mathbf{P}(\mathbf{r}, t))_{\ell\ell'} \int^t dt' (H_t^l(t'))_{\ell'\ell} \\
 &- \sum (i\hbar)^{-1} \sum_{\ell'} (\mathbf{P}(\mathbf{r}, t))_{\ell'\ell} \int^t dt' (H_t^l(t'))_{\ell\ell'}^* \quad (6.1.18)
 \end{aligned}$$

where we have made use of the interaction representation of $\mathbf{P}(\mathbf{r})$:

$$\mathbf{P}(\mathbf{r}, t) = \exp(itH/\hbar)\mathbf{P}(\mathbf{r})\exp(-itH/\hbar). \quad (6.1.19)$$

Although the macroscopic relation $\mathbf{J} = \partial\mathbf{P}/\partial t$ does not hold for operators themselves, one can readily prove the following equation in the spatially integrated form:

$$- \int d\mathbf{r}' \mathbf{J}(\mathbf{r}', t') \cdot \mathbf{A}(\mathbf{r}', t') = - \int d\mathbf{r}' \partial\mathbf{P}(\mathbf{r}', t')/\partial t' \cdot \mathbf{A}(\mathbf{r}', t'). \quad (6.1.20)$$

In fact, the Heisenberg equation of motion gives

$$\begin{aligned}
 \partial P_\alpha(\mathbf{r}', t')/\partial t' &= (i/\hbar) \sum_i (e_i/2m_i) \\
 &\times \left\{ \exp(it'H/\hbar) \sum_{\alpha'} [p_{i\alpha'}^2, r_{i\alpha} \Lambda(\mathbf{r}' - \mathbf{r}_i)] \exp(-it'H/\hbar) \right\} \quad (6.1.21)
 \end{aligned}$$

since only the kinetic energy contributes to the commutator with $\mathbf{P}(\mathbf{r})$ which contains the position coordinates \mathbf{r}_i . With the use of the Heisenberg representations of individual operators, one can rewrite the expression $\{\cdot\cdot\cdot\}$ in (6.1.21) as

$$(i/\hbar)[,] = p_{i\alpha'} \{ \delta_{\alpha\alpha'} \Lambda(\mathbf{r}' - \mathbf{r}_i) - r_{i\alpha} (\partial/\partial r_{\alpha'}) \Lambda(\mathbf{r}' - \mathbf{r}_i) \} + (h.c.).$$

The integration on \mathbf{r}' of the gradient term in $\{\cdot\cdot\cdot\}$ multiplied by the much more slowly varying function $A_{\alpha'}(\mathbf{r}', t')$ vanishes because

$$\int d\mathbf{r}' (\partial/\partial r_{\alpha'}) \Lambda(\mathbf{r}' - \mathbf{r}_i) = 0 \quad (6.1.22)$$

as is obvious from the fact that $\Lambda(\mathbf{r}' - \mathbf{r}_i)$ vanishes when $|\mathbf{r}' - \mathbf{r}_i|$ is greater than a_M (but still much smaller than λ). With the use of (6.1.15), one can thus bring the r.h.s. of eq. (6.1.20) into the form of the l.h.s.

The time integral in the second line of (6.1.18) can then be rewritten, by partial integration, as

$$\int^t dt' (H_t^l(t'))_{\ell'\ell} = - \int d\mathbf{r}' \int^t dt' \mathbf{J}(\mathbf{r}', t')_{\ell'\ell} \cdot \mathbf{A}(\mathbf{r}', t')$$

$$\begin{aligned}
&= - \int d\mathbf{r}' \int^t dt' \partial \mathbf{P}(\mathbf{r}', t')_{\ell\ell} / \partial t' \cdot \mathbf{A}(\mathbf{r}', t') \\
&= - \int d\mathbf{r}' \{ \mathbf{P}(\mathbf{r}', t)_{\ell\ell} \cdot \mathbf{A}(\mathbf{r}', t) - \int^t dt' \mathbf{P}(\mathbf{r}', t')_{\ell\ell} \cdot \mathbf{E}(\mathbf{r}', t') \}. \quad (6.1.23)
\end{aligned}$$

The external field $\mathbf{A}(\mathbf{r}', t')$ is assumed to vanish at $t' = -\infty$, the lower bound of integration so far not written explicitly, in order to assure the convergence of the integral. This assumption justifies the above omission of the integrand at the lower end.

In this way, the field induced part (the second and third lines of (6.1.18)) of the polarization is given by

$$\begin{aligned}
&\int d\mathbf{r}' \{ (i/\hbar) ([\mathbf{P}(\mathbf{r}, t), \mathbf{P}(\mathbf{r}', t) \cdot]_{-})_{\ell\ell} \mathbf{A}(\mathbf{r}', t) \\
&\quad + \int^t dt' (i/\hbar) ([\mathbf{P}(\mathbf{r}, t), \mathbf{P}(\mathbf{r}', t') \cdot]_{-})_{\ell\ell} \mathbf{E}(\mathbf{r}', t') \}. \quad (6.1.24)
\end{aligned}$$

Since $\mathbf{P}(\mathbf{r})$ contains only the position coordinates \mathbf{r}_i of the particles, the commutator $[\mathbf{P}(\mathbf{r}), \mathbf{P}(\mathbf{r}')]_{-}$ vanishes, and so does the first commutator in (6.1.24) which is its Heisenberg representation. Taking the statistical average on the initial states ℓ with Boltzmann distribution: $w_{\ell} = \exp[-\beta(E_{\ell} - F)]$ with free energy $F = -\beta^{-1} \ln[\sum_{\ell} \exp(-\beta E_{\ell})]$, one obtains the expression for the ij - (hereafter used for $\alpha\alpha'$ -) component of the susceptibility tensor χ defined in (6.1.1) as

$$\begin{aligned}
\chi_{ij}(\mathbf{r}; \mathbf{r}', t - t') &= \langle \langle (i/\hbar) [P_i(\mathbf{r}, t), P_j(\mathbf{r}', t')]_{-} \rangle \rangle \\
&= (i/\hbar) \sum_{\ell\ell'} w_{\ell} [P_i(\mathbf{r})_{\ell\ell'} P_j(\mathbf{r}')_{\ell'\ell} \exp\{-i\omega_{\ell\ell'}(t - t')\} \\
&\quad - P_j(\mathbf{r}')_{\ell\ell'} P_i(\mathbf{r})_{\ell'\ell} \exp\{i\omega_{\ell\ell'}(t - t')\}]. \quad (6.1.25)
\end{aligned}$$

The correlation functions on the r.h.s., such as

$$\sum_{\ell\ell'} w_{\ell} P_i(\mathbf{r})_{\ell\ell'} P_j(\mathbf{r}')_{\ell'\ell} = \langle \langle P_i(\mathbf{r}) P_j(\mathbf{r}') \rangle \rangle,$$

are expected to depend only on the difference $\mathbf{r} - \mathbf{r}'$ in homogeneous systems in agreement with (6.1.4). It is also important to note that the integration on t' , the time when the external field acts upon the system, has been extended up to t (and not beyond), the time when the response is recognized, without any definition of (6.1.25) at $t < t'$. This is consistent with the causality (6.1.2).

For practical applications, it may be more convenient to transform the description in spacetime into that in the reciprocal space, \mathbf{k} and ω . For this purpose, let us consider a volume $V = L^3$ with cyclic boundary conditions as in Section 1.3 and

introduce the Fourier transform:

$$\mathbf{E}(\mathbf{r}, t) = V^{-1/2} \sum_{\mathbf{k}} \int d\omega \mathbf{E}_{\mathbf{k}, \omega} \exp(i\mathbf{k} \cdot \mathbf{r} - i\omega t), \quad (6.1.26)$$

$$\mathbf{P}(\mathbf{r}, t) = V^{-1/2} \sum_{\mathbf{k}} \int d\omega \mathbf{P}_{\mathbf{k}, \omega} \exp(i\mathbf{k} \cdot \mathbf{r} - i\omega t), \quad (6.1.27)$$

$$\mathbf{P}(\mathbf{r}) = V^{-1/2} \sum_{\mathbf{k}} \mathbf{P}_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r}) \quad (\text{q-number}). \quad (6.1.28)$$

Making use of the homogeneity approximation (6.1.4) and rewriting $\chi(\mathbf{r} - \mathbf{r}', t - t') = \chi(\mathbf{s}, \tau)$, one can readily derive the susceptibility tensor $\chi(\mathbf{k}, \omega)$:

$$\mathbf{P}_{\mathbf{k}, \omega} = \chi(\mathbf{k}, \omega) \mathbf{E}_{\mathbf{k}, \omega}, \quad (6.1.29)$$

$$\chi(\mathbf{k}, \omega) = \lim_{\gamma \rightarrow +0} \int ds \int_0^\infty d\tau \chi(\mathbf{s}, \tau) \exp(-i\mathbf{k} \cdot \mathbf{s} + i\omega\tau - \gamma\tau) \quad (6.1.30)$$

where ω has been replaced by $\omega + i\gamma$ ($\gamma \rightarrow +0$) to consider the afore-mentioned initial condition that the external field (6.1.26) has been switched on infinitely slowly, which assures the convergence of τ -integration in (6.1.30). Putting (6.1.25) into (6.1.30), one obtains the susceptibility tensor:

$$\chi_{ij}(\mathbf{k}, \omega) = \sum_{\ell\ell'} w_\ell [(P_{i,\mathbf{k}})_{\ell\ell'} (P_{j,-\mathbf{k}})_{\ell'\ell} / (\hbar\omega_{\ell'\ell} - \hbar\omega - i\gamma) + (P_{j,-\mathbf{k}})_{\ell\ell'} (P_{i,\mathbf{k}})_{\ell'\ell} / (\hbar\omega_{\ell'\ell} + \hbar\omega + i\gamma)]. \quad (6.1.31)$$

One can readily confirm that this tensor has the properties:

$$\chi_{ij}(\mathbf{k}, \omega) = \chi_{ji}^*(\mathbf{k}, \omega) = \chi_{ji}(-\mathbf{k}, -\omega). \quad (6.1.32)$$

For those systems consisting of mobile (instead of bound) charges, it is more convenient to rewrite eq. (6.1.31) in terms of charge density $\rho(\mathbf{r}) = \text{div} \mathbf{P}(\mathbf{r})$, namely $\rho_{\mathbf{k}} = i\mathbf{k} \cdot \mathbf{P}_{\mathbf{k}}$, as

$$\chi(\mathbf{k}, \omega)_{\parallel} = k^{-2} \sum_{\ell\ell'} w_\ell [|(\rho_{-\mathbf{k}})_{\ell'\ell}|^2 / (\hbar\omega_{\ell'\ell} - \hbar\omega - i\gamma) + |(\rho_{\mathbf{k}})_{\ell'\ell}|^2 / (\hbar\omega_{\ell'\ell} + \hbar\omega + i\gamma)]. \quad (6.1.33)$$

The imaginary part of (6.1.33) can be rewritten, after replacement $\ell \rightleftharpoons \ell'$ in the second line, as

$$\text{Im}[\chi(\mathbf{k}, \omega)_{\parallel}] = (2\hbar k^2)^{-1} [1 - \exp(-\beta\hbar\omega)] \int_{-\infty}^{+\infty} dt \exp(-i\omega t) \langle \rho_{\mathbf{k}} \rho_{-\mathbf{k}}(t) \rangle. \quad (6.1.34)$$

For an isotropic system and at $\mathbf{k} = 0$, eq. (6.1.31) can be written as

$$\chi(0, \omega) = \chi_{ii}(0, \omega) = \sum_{\ell\ell'} w_\ell |(P_i)_{\ell\ell'}|^2 \times [(\hbar\omega_{\ell'\ell} - \hbar\omega - i\gamma)^{-1} + (\hbar\omega_{\ell'\ell} + \hbar\omega + i\gamma)^{-1}]. \quad (6.1.31a)$$

6.2 Analytic properties, dispersion relation and sum rule

It is obvious from (6.1.31) that the susceptibility tensor is an analytic function of the variable ω in the upper half of its complex plane, since all poles of (6.1.31) are on the lower half. This is the general property of all kinds of linear response coefficient, and is intimately related with the causality and the irreversibility of systems that they should approach thermal equilibrium when left alone.

Making use of this analyticity, one can perform the contour integral of $\chi_{ij}(\mathbf{k}, \omega')/(\omega' - \omega)$ with real ω along the path of ω' shown in Fig. 6.1, with the result:

$$\chi_{ij}(\mathbf{k}, \omega) = (i\pi)^{-1} \mathcal{P} \int_{-\infty}^{+\infty} d\omega' (\omega' - \omega)^{-1} \chi_{ij}(\mathbf{k}, \omega'). \quad (6.2.1)$$

Taking the real and imaginary parts of (6.2.1), one obtains the dispersion relations:

$$\text{Re} \chi_{ij}(\mathbf{k}, \omega) = \pi^{-1} \mathcal{P} \int_{-\infty}^{+\infty} d\omega' (\omega' - \omega)^{-1} \text{Im} \chi_{ij}(\mathbf{k}, \omega'), \quad (6.2.2)$$

$$\text{Im} \chi_{ij}(\mathbf{k}, \omega) = -\pi^{-1} \mathcal{P} \int_{-\infty}^{+\infty} d\omega' (\omega' - \omega)^{-1} \text{Re} \chi_{ij}(\mathbf{k}, \omega'). \quad (6.2.3)$$

By putting $\mathbf{k} = 0$ in the imaginary part of (6.1.31), and interchanging ℓ and ℓ' in the second term, one obtains

$$\text{Im} \chi_{ij}(0, \omega) = (\pi/h) \sum_{\ell\ell'} (w_\ell - w_{\ell'}) (P_{i,0})_{\ell\ell'} (P_{j,0})_{\ell'\ell} \delta(\omega_{\ell'\ell} - \omega), \quad (6.2.4)$$

from which one finds that $\text{Im} \chi_{ij}(0, \omega)$ is an odd function of ω . Applying the sum rule which was derived in Section 3.6, one can derive the related sum rule for an assembly of charged particles with mass m_i and electric charge e_i within a unit volume:

$$\int_{-\infty}^{+\infty} \omega \text{Im} \chi_{ij}(0, \omega) d\omega = \int_{-\infty}^{+\infty} \text{Re} \chi_{ij}(0, \omega) d\omega = \delta_{ij} \pi \sum_i e_i^2 / m_i \quad (6.2.5)$$

with the use of the equations

$$\sigma \mathbf{E} = \mathbf{J} = d\mathbf{P}/dt = -i\omega \mathbf{P}, \quad \text{or} \quad \sigma(\mathbf{k}, \omega) = -i\omega \chi(\mathbf{k}, \omega) \quad (6.2.6)$$

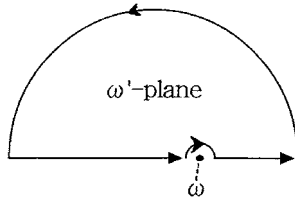


Fig. 6.1 The path for the integration on the complex ω' -plane for the derivation of the dispersion formula.

and eq. (6.1.6). Equation (6.2.5) will have various applications in later sections since the integrand with $i = j$ is related to the absorption spectra.

6.3 Relation between electric susceptibility and dielectric constant

In Sections 6.1 and 6.2 we derived the microscopic expressions for electric susceptibility χ and conductivity σ (see eq. (6.2.6) for their relationship) by assuming that $\mathbf{E}(\mathbf{r})$ acting on the charged particles of the system is an *external* field (hereafter written $\mathbf{E}_{\text{eff}}(\mathbf{r})$). Note, however, that the electric field $\mathbf{E}(\mathbf{r})$ appearing in the Maxwell equations and in definition (1.1.5) of the dielectric constant ϵ is a *macroscopic* quantity, namely the spatial average over the region of macroscopic size \mathbf{r} , of the microscopic electric field $\mathbf{e}(\mathbf{r})$ which fluctuates in the microscopic region. Now, $\mathbf{e}(\mathbf{r})$ consists of the effective internal field $\mathbf{E}_{\text{eff}}(\mathbf{r})$ and the internal field due to the charged particles in the system which is responsible for the microscopic fluctuation. Thus, $\mathbf{E}_{\text{eff}}(\mathbf{r})$, which is the macroscopic average of $\mathbf{e}_{\text{eff}}(\mathbf{r})$, is *in general different* from the macroscopic $\mathbf{E}(\mathbf{r})$. Microscopic expressions for $\mathbf{e}_{\text{eff}}(\mathbf{r})$ and $\mathbf{e}(\mathbf{r})$ as well as the relation between susceptibility and dielectric constant, *valid for general systems*, have unfortunately *not* been given so far. Here we will put in order different ways of relating them in terms of two approaches, the \mathbf{E} method and the \mathbf{D} method,³ taking typical model systems – insulators and conductors.

Under no external field, $\mathbf{e}(\mathbf{r}) \equiv -\nabla\phi_{\text{micro}}(\mathbf{r})$ consists of the electric field due to all charged particles of the system considered. If the ions are arrayed on a crystalline lattice, they will give rise to a periodic potential $\phi_{\text{micro}}(\mathbf{r})$. If, in addition, there are free electrons, as in metals, their time-averaged density will also contribute to this periodic potential. The individual free electrons are subject to the potential $(-e)\phi_{\text{micro}}(\mathbf{r})$, forming the conduction band with effective mass m_c . Obviously, the average of $\mathbf{e}(\mathbf{r})$ over the macroscopic region vanishes. The remaining parts of the interactions yet to be taken into account are the temporally fluctuating parts of the electron–electron, ion–ion and electron–ion interactions which are supposed to be described by the screened Coulomb interactions (see Ref. [4] or Section 6.5 of the present book), phonon energies (after including the kinetic energies of the ions not mentioned above) and the electron–phonon interactions (described in Sections 9.1 and 9.4), respectively. They altogether will be denoted by an effective Hamiltonian: H_{eff} .

Application of an electric field will induce polarization, namely the redistribution of electrons and ions which, in turn, gives rise to a depolarizing field as was already mentioned in Section 5.3 for lattice vibrations in an insulator. The redistribution of free electrons in conductors results in surface charges in the regions adjacent to the electrodes, thus screening the field $\epsilon_0^{-1}\mathbf{D}(\mathbf{r})$ originating from the electric charges on

the electrodes to $\mathbf{E}(\mathbf{r})$ in the metal, if averaged over the macroscopic region. The use of this $\mathbf{E}(\mathbf{r})$ for the effective field $\mathbf{E}_{\text{eff}}(\mathbf{r})$ acting on the free electrons is well justified since their de Broglie wavelengths are usually much greater than the interatomic distance. As for insulators, we used $\mathbf{E}(\mathbf{r})$ for $\mathbf{E}_{\text{eff}}(\mathbf{r})$ in step (II) of Section 5.3 to calculate χ and ϵ (see eq. (5.3.7)) which are related by

$$\epsilon(\mathbf{k}, \omega) - \epsilon_0 = (\epsilon_e - \epsilon_0) + \chi_{\text{eff}}(\mathbf{k}, \omega). \quad (6.3.1)$$

Here we have phenomenologically separated the instantaneous susceptibility $(\epsilon_e - \epsilon_0)$ contributed from oscillators with much higher frequencies than ω . The Hamiltonians to be used for the calculation of χ for insulator and σ for conductors are H_{eff} mentioned above, for instance, the use of short-ranged \overline{f} instead of the original force constants f in insulators (as was in fact done in Section 5.3) and the use of short-ranged electron–electron interactions in conductors (as will be described in Section 6.5).

The above procedure of putting $\mathbf{E}_{\text{eff}}(\mathbf{r}) \simeq \mathbf{E}(\mathbf{r})$ in calculating χ and σ by (6.1.31) and (6.2.6) and relating χ to ϵ by (6.3.1) will hereafter be called the \mathbf{E} method. It can be applied to a wide class of systems inclusive of anisotropic systems, provided that one can find an appropriate or approximate H_{eff} . Complete neglect of Coulomb interactions among electrons in calculating the electric conductivity of a metal is an example of this method at its extreme simplification (even the short-ranged interaction remaining in H_{eff} is neglected there); in spite of its apparent crudity and the skeptical view on its legitimacy in its early days (in fact people wondered why the free-electron model for metal due to Sommerfeld worked so well), it has proved quite successful in explaining the observed temperature dependence of electric resistivity in terms of electron–phonon interaction alone.

In general, however, it is difficult to find such an H_{eff} on a sound theoretical basis since there is no unique way of decomposing the Coulomb interaction into long- and short-range parts. This ambiguity is a demerit of the \mathbf{E} method.

Another important situation in which the \mathbf{E} method is useful is an *isotropic* system under an electromagnetic wave. Because of the absence of the depolarizing field in the transverse wave, one is allowed to take \mathbf{E} approximately as the *external* field and to use a *real* Hamiltonian in calculating the *transverse* susceptibility $\chi_t(k, \omega)$ by the formula (6.1.31). Its relation with the dielectric constant is the same as (6.3.1):

$$\epsilon(k, \omega) - \epsilon_e = \chi_t(k, \omega) \quad (6.3.2)$$

where \mathbf{k} is replaced by k since the dependence on the direction \mathbf{k} is also excluded by our definition. In Section 8.4 we will calculate the energies of transverse excitons and the corresponding dielectric dispersion $\epsilon(\omega)$ with the use of the *real* Hamiltonian with Coulomb interactions fully considered, and the energies of longitudinal

excitons will be given by the zeros of thus obtained $\epsilon(\omega)$ in the same way as we prescribed for the optical lattice vibrations in Section 5.3.

For the study of eigenvibrations of a system under no external field, one has to add the self-induced depolarizing field (long range) to the short-range restoring force. This was done in step (III) of Section 5.3 for optical lattice vibrations, where the depolarizing field causes ℓ -t splitting. In conductors, the depolarizing field due to the redistribution of free electrons is the only restoring force for plasma oscillations (see Section 6.5). A more systematic way of describing the effect of the depolarizing field is the **D** method.

When the external field is longitudinal as is caused by charge density $\rho^{(0)}$ of incident or extrinsic charged particles, one must resort to the **D** method which is useful only for isotropic systems. The logic of this **D** method is as follows. The charge density, $\rho^{(0)}$, gives rise to an electric flux density **D** through the relation $\text{div} \mathbf{D} = \rho^{(0)}$ (eq. (1.1.3)), namely $i\mathbf{k} \cdot \mathbf{D}_k \equiv ikD_{\parallel} = \rho_k^{(0)}$, and hence to the electric field $\mathbf{E}^{(0)}$ through the relation $\mathbf{D} = \epsilon_e \mathbf{E}^{(0)}$, namely $E_{\parallel}^{(0)} = \epsilon_e^{-1} D_{\parallel}$. In turn $E_{\parallel}^{(0)}$, as the external field, induces electric polarization through the relation $P_{\parallel} = \chi_{\ell} E_{\parallel}^{(0)}$ where $\chi_{\ell}(k, \omega)$ is the *longitudinal* susceptibility. The internal electric field **E** is given by the identity $\mathbf{D} - \mathbf{P} = \epsilon_e \mathbf{E}$. Dividing both sides by **D**, one obtains

$$1 - \epsilon_e/\epsilon(k, \omega) = \chi_{\ell}(k, \omega)/\epsilon_e \quad (\text{isotropic, use of } H). \quad (6.3.3)$$

One must make use of the *real* Hamiltonian in calculating $\chi_{\ell}(k, \omega)$ by (6.1.31) or (6.1.33) since $\mathbf{E}^{(0)}$ is the purely external field.

So long as the macroscopic fields, $\mathbf{E}(\mathbf{k}, \omega)$ and $\mathbf{D}(\mathbf{k}, \omega)$, are definable in dielectrics and their linear relation holds, the proportionality coefficient $\epsilon(\mathbf{k}, \omega)$ exists independent of the way of relating it to the microscopically calculated susceptibility. Eliminating this $\epsilon(\mathbf{k}, \omega)$ from (6.3.1) or (6.3.2) and (6.3.3), one can relate the two kinds of susceptibility by

$$\chi_{\ell}/\epsilon_e = \chi_t/(\epsilon_e + \chi_t), \quad \chi_t/\epsilon_e = \chi_{\ell}/(\epsilon_e - \chi_{\ell}) \quad (6.3.4)$$

where χ_t can also be replaced by χ_{eff} of the isotropic case.

The above relations can most readily be confirmed with lattice vibrations in an isotropic crystal. In fact one finds from (5.3.13) the expressions

$$\chi_t/\epsilon_e = \Pi_a(\omega_{at}^2 - \omega^2)/(\omega_{at}^2 - \omega^2) - 1, \quad (6.3.5)$$

$$\chi_{\ell}/\epsilon_e = 1 - \Pi_a(\omega_{at}^2 - \omega^2)/(\omega_{at}^2 - \omega^2), \quad (6.3.5a)$$

with poles at transverse or longitudinal eigenfrequencies, as they should be.

Equation (6.3.4) has other applications. If one can find, by an appropriate physical argument on a given many-body system, an effective Hamiltonian H_{eff} which allows easy calculation of χ_{eff} with the use of the **E** method, one readily

obtains, with the use of (6.3.4), a first approximation to χ_ℓ which would be obtained only with a complicated diagrammatic method by direct application of the D method. An example is the plasma oscillation in an electron gas described in Section 6.5.

The E method is a type of the so-called molecular field theory, mean field theory or random phase approximation which has proved to be a useful first approximation to a variety of many-body problems in condensed matter and nuclear physics.

6.4 Dielectric constant in terms of phonons

So far we have described the contribution of optical lattice vibrations to the dielectric constant by classical mechanics. We will show here that quantum mechanics leads to exactly the same result. It is another example of the classical–quantal equivalence of harmonic oscillators.

Equation (5.5.2) is a quantum-mechanical expression for the displacement of the ν th atom in the n th unit cell:

$$\xi_{\nu n} = \sum_{s,k} (\hbar/2N\omega_{s,k})^{1/2} M_\nu^{-1/2} \zeta_\nu^{(s)}(\mathbf{k}) \times [b_{s,k} \exp(i\mathbf{k} \cdot \mathbf{R}_n) + b_{s,k}^\dagger \exp(-i\mathbf{k} \cdot \mathbf{R}_n)] \quad (6.4.1)$$

where $\zeta_\nu^{(s)}(\mathbf{k})$ is the normalized eigenvector of the (s, \mathbf{k}) mode of the lattice vibration with frequency $\omega_{s,k}$ in the *real* system (see Section 5.2). Comparing (5.3.6) with (6.1.27), one can write the Fourier component of polarization for small values of \mathbf{k} as

$$P_{\mathbf{k}} = \sum_s (\hbar N_0/2\omega_s)^{1/2} \mathbf{q}^{(s)}(b_{s,\mathbf{k}} + b_{s,-\mathbf{k}}^\dagger) \quad (6.4.2)$$

where N_0 is the number of unit cells within a unit volume and

$$\mathbf{q}^{(s)} \equiv \sum_\nu e_\nu M_\nu^{-1/2} \zeta_\nu^{(s)} \quad (6.4.3)$$

is a quantity corresponding to $\mathbf{p}^{(s)}$ of (5.3.4) for a fictitious system. Noting that $b_{s,\mathbf{k}}^\dagger$ and $b_{s,\mathbf{k}}$ have matrix elements (2.2.23) for creating or annihilating one phonon of the mode (s, \mathbf{k}) with energy change $\omega_{\ell'\ell} = \pm\Omega_s$, and that $\sum_{\ell'} [\dots]$ in (6.1.31) gives the ℓ -independent term as the result of summation on $\ell' = \ell \pm 1$, one finally obtains the temperature-independent result

$$\chi(\omega) = N_0 \sum_s \mathbf{q}^{(s)} \mathbf{q}^{(s)} / (\omega_s^2 - \omega^2), \quad (6.4.4)$$

which is the same as would have been obtained by classical mechanics. (In fact the corresponding term in (5.3.7) for the fictitious term was obtained by classical

mechanics.) Here we find again the classical–quantal equivalence as we did in (2.2.20), which is characteristic of harmonic oscillations. Confining ourselves to an isotropic system, one can put (6.4.4) into (6.3.3) to obtain $\epsilon(k, \omega)$, which should be the same as that $\epsilon(k, \omega)$ which appears in (6.3.1), as was confirmed in the last section.

6.5 Plasma oscillation and screening effect of electron gas

Consider an electron gas in a background of (almost) immobile neutralizing positive charges. Examples are carriers in the conduction band of metals, those in the conduction or valence band of highly doped and/or high-temperature semiconductors and an ionized gas at high temperature or low density. Here Coulomb interactions play central roles, resulting in collective oscillations and charge screening, as has been described by well-known methods for many-body problems.⁴

An alternative approach to the same problem will be presented here. Let us consider, as a first approximation, the effect of Coulomb interactions only through the macroscopic electric field, forgetting the individual interactions completely. This is the simplest example of the \mathbf{E} method mentioned in Section 6.3.

The one-particle state \mathbf{k} with energy $\varepsilon(\mathbf{k})$ is occupied with the Fermi distribution: $f(\mathbf{k}) = [\exp\{\beta(\varepsilon(\mathbf{k}) - \mu)\} + 1]^{-1}$ where μ is the chemical potential determined by the number of electrons, n , per unit volume, as

$$\sum_{\mathbf{k}} 2f(\mathbf{k}) = n \quad (6.5.1)$$

where the spin degeneracy 2 is taken into account. Consider the charge density operator defined by

$$\rho(\mathbf{r}) = (-e) \sum_i \delta(\mathbf{r}_i - \mathbf{r}), \quad (6.5.2)$$

or its Fourier component defined as in (6.1.26):

$$\rho_{-\mathbf{q}} = V^{-1/2} \sum_i \exp(i\mathbf{q} \cdot \mathbf{r}_i) \quad (6.5.3)$$

where \mathbf{q} for the density wave is used to avoid confusion with \mathbf{k} for the individual electrons. The operator (6.5.3), which appears in (6.1.33) with \mathbf{q} being written as \mathbf{k} , has matrix element which corresponds to one electron, say the i th, being scattered from state \mathbf{k} to $\mathbf{k} + \mathbf{q}$. Since the probability for the final state to be vacant and hence available is given by $1 - f(\mathbf{k} + \mathbf{q})$, the sum of the first term in (6.1.33) with spin degeneracy considered can be written as

$$V^{-1} \sum_{\mathbf{k}} e^2 2f(\mathbf{k}) [1 - f(\mathbf{k} + \mathbf{q})] / [\varepsilon(\mathbf{k} + \mathbf{q}) - \varepsilon(\mathbf{k}) - \hbar\omega].$$

The second term is obtained by changing the signs of \mathbf{q} and ω , but with further replacement $\mathbf{k} \rightarrow -\mathbf{k}$ or $\mathbf{k} \rightarrow \mathbf{k} + \mathbf{q}$, one obtains the final results:

$$\chi_{\text{eff}}(\mathbf{q}, \omega) = (e^2/q^2)V^{-1} \sum_{\mathbf{k}} 4f(\mathbf{k})[\varepsilon(\mathbf{k}) - \varepsilon(\mathbf{k} + \mathbf{q})]/[(\hbar\omega)^2 - \{\varepsilon(\mathbf{k}) - \varepsilon(\mathbf{k} + \mathbf{q})\}^2] \quad (6.5.4)$$

$$= (e^2/q^2)V^{-1} \sum_{\mathbf{k}} 2f(\mathbf{k})[f(\mathbf{k}) - f(\mathbf{k} + \mathbf{q})]/[\varepsilon(\mathbf{k} + \mathbf{q}) - \varepsilon(\mathbf{k}) - \hbar\omega]. \quad (6.5.4a)$$

Putting $\varepsilon(\mathbf{k}) = \hbar^2 k^2/2m$ with isotropic effective mass m of the electrons, one find α_{eff} also to be isotropic. In the vanishing limit of q , eq. (6.5.4) gives the temperature-independent value:

$$\chi_{\text{eff}}(\omega) = -ne^2/m\omega^2 \quad (6.5.5)$$

where (6.5.1) has been used. From (6.3.4), one obtains

$$\chi(\omega) = (ne^2/m)/(\omega_p^2 - \omega^2), \quad (6.5.6)$$

$$\omega_p^2 = ne^2/\epsilon_e m \quad (6.5.7)$$

where ϵ_e is the background dielectric constant due to bound electrons. Equation (6.5.6) indicates that the electron gas has eigenvibration with angular frequency ω_p . As is obvious from (6.5.6) and (6.5.7), this vibration is the density oscillation of the electron gas caused entirely by the depolarizing field. This collective motion of the electron gas is called the *plasma* oscillation. In the harmonic approximation, the plasma oscillation field can be described quantum-mechanically as an assemblage of plasmons with energy quanta $\hbar\omega_p$, analogous to the photons of an electromagnetic field and the phonons of lattice vibrations.

Let us next put $\omega = 0$ and q to be small in (6.5.4a). Noting that the Fermi distribution $f(\mathbf{k})$ is a function of $\varepsilon = \varepsilon(\mathbf{k})$, one finds that the summand can be written as $-2\partial f(\varepsilon)/\partial \varepsilon$. The static dielectric constant is then given by

$$\epsilon(q, 0) = \epsilon_e(1 + q_s^2/q^2), \quad (6.5.8)$$

$$\text{where } q_s^2 = (e^2/\epsilon_e)V^{-1} \sum_{\mathbf{k}} [-2\partial f(\varepsilon)/\partial \varepsilon]. \quad (6.5.9)$$

For the two limiting situations of statistics, one obtains

$$q_s^2 = ne^2/\epsilon_e(2\varepsilon_F/3) \quad (\text{degenerate Fermi distribution}), \quad (6.5.10)$$

$$q_s^2 = ne^2/\epsilon_e k_B T \quad (\text{Boltzmann distribution}), \quad (6.5.10a)$$

since one can replace $-\partial f(\varepsilon)/\partial \varepsilon$ by $\delta(\varepsilon - \varepsilon_F)$ and β in the respective cases. Here the Fermi energy $\varepsilon_F = \hbar^2 k_F^2/2m$ is the energy of the highest occupied states: $k_F = (3\pi^2 n)^{1/3}$ in the degenerate case.

The physical meaning of q_s is as follows. Suppose a point charge Q is put at $\mathbf{r} = 0$. From $\text{div} \mathbf{D}(\mathbf{r}) = Q\delta(\mathbf{r})$, we have the Fourier component $\mathbf{D}_q = V^{-1/2} Q\mathbf{q}/iq^2$ and that of the electric field $\mathbf{E}_q = \mathbf{D}_q/\epsilon(\mathbf{q}, 0)$. From $\mathbf{E}(\mathbf{r}) = -\text{grad}\phi(\mathbf{r})$, we have the Fourier component of the electrostatic potential $\phi_q = V^{-1/2} Q/q^2\epsilon(\mathbf{q}, 0)$. With the use of (6.5.8), we finally obtain

$$\phi(\mathbf{r}) = V^{-1/2} \sum_q \phi_q \exp(i\mathbf{q} \cdot \mathbf{r}) = (Q/4\pi\epsilon_c r) \exp(-q_s r). \quad (6.5.11)$$

It indicates that the electrostatic potential due to the point charge $Q/4\pi\epsilon_0 r$ is (dielectrically) screened to $Q/4\pi\epsilon_c r$ by the bound electrons and (metallically) screened to the short-range potential (6.5.11) with *screening distance* q_s^{-1} by the mobile electrons. The screening constant q_s is given by (6.5.10) or (6.5.10a) depending upon the situation.

6.6 Energy loss of a charged particle injected into a dielectric

Suppose a particle with charge Ze and momentum $\hbar\mathbf{K}$ is injected into the material system. It gives rise to the electrostatic potential $\phi(\mathbf{r} - \mathbf{R}) = Ze/4\pi\epsilon_0|\mathbf{r} - \mathbf{R}|$ which interacts with the system through the Hamiltonian

$$H' = \int d\mathbf{r} \rho(\mathbf{r})\phi(\mathbf{r} - \mathbf{R}) = \sum_k \phi_k \rho_{-k} \exp(-i\mathbf{k} \cdot \mathbf{R}). \quad (6.6.1)$$

Here $\rho(\mathbf{r})$ is the charge-density operator of the system and $\phi_k = V^{-1/2} Ze/\epsilon_0 k^2$ is the Fourier component of $\phi(\mathbf{r} - \mathbf{R})$. The rate at which this particle loses momentum $\hbar\mathbf{k}$ and hence energy $\hbar\omega = \hbar^2\mathbf{K}^2/2M - \hbar^2(\mathbf{K} - \mathbf{k})^2/2M$ by exciting the system in a state ℓ to a state ℓ' is given, in the Born approximation which is valid for high-energy particles, by

$$W(k, \omega) = (2\pi/\hbar) |\phi_k|^2 \sum_{\ell\ell'} w_{\ell} |(\rho_{-\mathbf{k}})_{\ell'\ell}|^2 \delta(\hbar\omega_{\ell'\ell} - \hbar\omega). \quad (6.6.2)$$

While this can be related to the first term of (6.1.33), the susceptibility is to be related to the dielectric constant by (6.3.3), namely by the \mathbf{D} method since the external force is due to the charged particle. Hence, the rate can be written in terms of the dielectric constant as

$$W(k, \omega) = (2/\hbar) |\phi_k|^2 k^2 \text{Im}[-1/\epsilon_{\parallel}(\mathbf{k}, \omega)]. \quad (6.6.3)$$

To sum up, the injected charged particle excites the longitudinal component of polarization through $\text{Im}[-1/\epsilon_{\parallel}(\mathbf{k}, \omega)]$ while the electromagnetic wave excites the transverse component through $\text{Im}[\epsilon(\mathbf{k}, \omega)]$. The plasma, the charge-density oscillation, is obviously a longitudinal wave and can be excited by the charged particle. This is how the plasma oscillations have been observed through electron energy-loss spectra (EELS), (6.6.3), of high-energy charged particles.

6.7 Optical activity

As a typical application of the anisotropy and spacial dispersion of the susceptibility given by (6.1.31), we consider here the optical activity of a material in the simplest case. Optical activity refers to the phenomena in which the plane of polarization (containing \mathbf{E} and \mathbf{k} vectors) of linearly polarized light rotates gradually as the light beam proceeds⁵ through materials with low symmetry. In other words, the circularly polarized light beams, rotating right or clockwise (+) and left or anticlockwise (−) (as viewed by the observer faced towards the light source) propagate with different phase velocities due to different refractive indices, n_+ and n_- . It seems to be of particular significance in clarifying the origin of life since biomolecules on Earth consist of *isomers* that are either 100% right-handed or 100% left-handed depending upon the molecular species such as carbohydrates, proteins and nucleic acids. This remarkable fact, called *homochirality*, is one of the greatest wonders in life science. It is to be contrasted with the fact that, irrespective of the species, biomolecules synthesized by *abiotic* experiments in chemical laboratories always consist of equal amounts of right-handed and left-handed isomers. It is also known that a dilute solution of *identical isomers* in an *optically inactive solvent* shows optical activity opposite to that of *opposite isomers with a spacially inverted structure*, in spite of the fact that the solute molecules have random orientations in dilute solution. This indicates that the optical activity is already inherent to a molecule although there are a number of optically active crystals, such as quartz, which consist of both right- and left-handed isomers. For this reason we consider here a dilute solution of identical solute molecules in an optically inactive solvent which seems to be the simplest model in which to consider optical activity.

Let us first study the symmetry of the dielectric tensor $\bar{\epsilon}(\mathbf{k}, \omega)$ of such a solution with z -axis taken in the \mathbf{k} direction. From the macroscopic isotropy of the system, one has $\epsilon_{xx} = \epsilon_{yy}$ due to the invariance against rotation around the z -axis. From the general property of the susceptibility tensor, eq. (6.1.32), one has $\epsilon_{yx} = \epsilon_{xy}^*$. However, $\epsilon_{yx} = \epsilon_{xy}$ is not valid due to the absence of inversion symmetry (note that $x \rightarrow y, y \rightarrow x, z \rightarrow z$ is a mirror reflection which is not allowed).

One can calculate the susceptibility tensor $\bar{\chi}(\mathbf{k}, \omega) = \epsilon(\mathbf{k}, \omega) - \epsilon_0$ from (6.1.31) with the use of an appropriate microscopic model. We begin with the contribution of one solute molecule, with an electronic transition from the ground state g to an excited state e , to the *off*-diagonal element χ_{xy} which is essential for the optical activity. With the use of eqs. (6.1.6), (6.1.32) and the relation $\mathbf{J}_{eg} = (d\mathbf{P}/dt)_{eg} = i\omega_{eg}\mathbf{P}_{eg}$, one obtains the contribution of the molecule at \mathbf{R} to a matrix element appearing in (6.1.31):

$$(P_{y,-k})_{eg} = V^{-1/2}(e/im\omega_{eg})\exp(i\mathbf{k} \cdot \mathbf{R})[\{p_y \exp(ikz) + \exp(ikz)p_y\}/2]_{eg}. \quad (6.7.1)$$

The term $[\dots]_{eg}$ can be expanded in k as $(p_y + ikp_yz + \dots)_{eg}$. The corresponding matrix element $[\dots]_{ge}$ in $(P_{xk})_{ge}$ turns out to be $(p_x - ikp_{xz} + \dots)_{ge}$. Multiplying this with the above $[\dots]_{eg}$, we obtain a factor

$$(p_x)_{ge}(p_y)_{eg} + ik[(p_x)_{ge}(p_yz)_{eg} - (p_{xz})_{ge}(p_y)_{eg}] + \dots \quad (6.7.2)$$

Equation (6.7.2) has been calculated for a fixed orientation of the molecule under consideration.

Assuming that the solute molecules in the dilute solution take random orientations without any correlation among neighboring ones, one must take the average of the orientations with the use of the Eulerian angles, and multiply by N , the total number of solute molecules within the volume V . (In summing the contributions of all solvent molecules, we note that the position-dependent factor $\exp(i\mathbf{k} \cdot \mathbf{R})$ in (6.7.1) multiplied by $\exp(-i\mathbf{k} \cdot \mathbf{R})$ of $(P_{xk})_{ge}$ becomes position independent.) Only the second term of (6.7.2) remains, giving

$$N_0 \langle ([\dots]_{ge}[\dots]_{eg}) \rangle_{\text{orientation}} = -ikN_0 \text{Re}[(1/3) \sum_{xyz} (p_x)_{ge}(l_x)_{eg}] + \dots, \quad (6.7.3)$$

where $l_x \equiv yp_z - zp_y$ is the x -component of the angular momentum. Being of the same form as the scalar product, \sum can be written as $\mathbf{p}_{ge} \cdot \mathbf{l}_{eg}$ which is invariant against the choice of the coordinate system, as it should be, being a quantity *inherent to the molecule*. It is only for a molecule without inversion symmetry that the matrix elements of linear momentum \mathbf{p} and angular momentum \mathbf{l} can coexist for the same pair of states e and g , which is essential for the optical activity as will now be seen. Moreover, the nonvanishing *scalar* product (the component of \mathbf{l} along \mathbf{p}) is the reflection of a sort of helical structure which distinguishes right- and left-handed isomers.

As can be readily seen in the above argument, \mathbf{l} originates in the \mathbf{r} -dependent part of the response $\mathbf{P}(\mathbf{r})$ and hence of the external field $\mathbf{A}(\mathbf{r})$. As can be seen from (3.3.7), the interaction of an electron with the \mathbf{k} -component of electromagnetic field $\mathbf{A}_k(\mathbf{r})$ can be expanded as

$$H_I = (e/m)\mathbf{p} \cdot \mathbf{A}_k(\mathbf{r}) = (e/m)(\mathbf{p} \cdot \mathbf{e})A_0(1 + i\mathbf{k} \cdot \mathbf{r} + \dots) \quad (6.7.4)$$

where \mathbf{e} is the polarization vector of the electromagnetic field \mathbf{A}_k and hence is parallel to the electric field \mathbf{E}_k . The first term H_I' in the expansion on the r.h.s. of (6.7.4) gives rise to the optical transition associated with the *electric dipole moment* as is considered in the major part of this book and has appeared as $(\mathbf{p})_{ge}$ in (6.7.3). The second term can be rewritten, with the use of the well-known identity of vector algebra: $(\mathbf{k} \times \mathbf{e}) \cdot (\mathbf{r} \times \mathbf{p}) = (\mathbf{k} \cdot \mathbf{r})(\mathbf{e} \cdot \mathbf{p}) - (\mathbf{e} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{p})$,

as follows:

$$H_I'' = (ie/m)A_0(\mathbf{p} \cdot \mathbf{e})(\mathbf{k} \cdot \mathbf{r}) = (ie/2m)A_0[(\mathbf{e} \cdot \mathbf{p})(\mathbf{k} \cdot \mathbf{r}) + (\mathbf{e} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{p})] \\ + (ie/2m)A_0(\mathbf{k} \times \mathbf{e}) \cdot (\mathbf{r} \times \mathbf{p}). \quad (6.7.5)$$

Since $\mathbf{B}_0 = (i\mathbf{k} \times A_0\mathbf{e})$ gives the amplitude of magnetic flux density $\mathbf{B} = \nabla \times \mathbf{A}$ of the electromagnetic field and $\mathbf{r} \times \mathbf{p} \equiv \mathbf{l}$ is the orbital angular momentum, the second term on the r.h.s. of (6.7.5) gives rise to the optical transition associated with the *magnetic dipole moment* with the Landé factor $(e/2m)$. In the expression (6.7.3), the angular momentum \mathbf{l} appears in the form of a scalar product with linear momentum \mathbf{p} , indicating the role of helicity of the molecule in the optical activity, as mentioned above. The first term of (6.7.5) gives rise to that associated with the *electric quadrupole moment* which is irrelevant in the present problem of optical activity. We will see another example of the magnetic dipole playing a role in two-photon absorption in Section 11.2 (see Ref. [23] of Chapter 11).

The off-diagonal component of the electric susceptibility tensor (6.1.31) is then given by

$$\chi_{xy}(\mathbf{k}, \omega) = -\chi_{yx}(\mathbf{k}, \omega) \\ = (2N_0e^2/3\hbar m^2) \sum_e \omega_{eg}^{-1} \{\omega_{eg}^2 - (\omega + i\gamma)^2\}^{-1} [-ik \text{Re}(\mathbf{p}_{ge} \cdot \mathbf{l}_{eg})], \quad (6.7.6)$$

while the corresponding contribution to the diagonal element is given by replacing $[\cdot \cdot \cdot]$ in (6.7.6) by $(\mathbf{p}_{ge} \cdot \mathbf{p}_{eg})$.

Consider now the right and left circularly polarized light beams proceeding in the z-direction:

$$E^{(\pm)} = (E_x \pm iE_y) = E \exp(ik_{\pm}z - i\omega t). \quad (6.7.7)$$

The corresponding electric flux density is given by

$$D^{(\pm)} = (\epsilon_{xx} \mp i\chi_{xy})E^{(\pm)} = \epsilon_0 n_{\pm} E^{(\pm)}, \quad (6.7.8)$$

$$k_{\pm} = (\omega/c_0)n_{\pm} = (\omega/c)[1 \mp i\chi_{xy}/2\epsilon_{xx}]. \quad (6.7.9)$$

The polarization vector of a linearly polarized light beam, which can be expressed as a superposition of the right and left circularly polarized beams, rotates by an angle:

$$\phi = (k_- - k_+)z/2 = \pi(n_- - n_+)z/\lambda_0, \quad \lambda_0 \equiv c_0/\omega \quad (6.7.10)$$

for distance z . The rotatory power is therefore given by

$$\phi/z = \pi(\epsilon_0\epsilon_{xx})^{-1/2}(-i\chi_{xy}) = (2\pi^2/\epsilon_0\lambda_0^2)(2N_0e^2/3\hbar m^2) \\ \times \sum_e \omega_{eg}^{-1} \{\omega_{eg}^2 - (\omega + i\gamma)^2\}^{-1} \text{Re}(\mathbf{p}_{ge} \cdot \mathbf{l}_{eg}) \quad (6.7.11)$$

which, in the transparent region ($\omega < \omega_{eg}$), is of the order of $(N_0 a^3) a \lambda_0^{-2}$ where a is the molecular radius. Namely, the rotation angle per wavelength is of the order of concentration times a/λ_0 , which is quite small since typically $a \sim 10^{-9}\text{m}$ while $\lambda_0 \sim 6 \times 10^{-7}\text{ m}$ for visible light.

To conclude, it is emphasized again that the *general* theory for the dielectric tensor with *spatial dispersion* is enough for the derivation of a concrete *molecular* expression of optical activity with coexistent matrix elements of \mathbf{l} and \mathbf{p} indicative of helicity (nonvanishing only for molecules without parity), the fact which indicates the usefulness of the general theory. The situation resembles the helicity of an elementary particle defined by $h = \mathbf{s} \cdot \mathbf{p} / |\mathbf{p}|$ where s is the spin instead of the orbital angular momentum. For instance, neutrinos and anti-neutrinos have helicity of opposite signs, $-1/2$ (left-handed screw) and $+1/2$ (right-handed screw), respectively, as has been clarified in the famous experiment confirming the violation of parity.

One-electron states in solids

7.1 Energy band

In this chapter, we shall describe very briefly the electronic states of crystals, based on a one-electron picture and with the tight binding approximation, mainly as a basis for the many-body problems to be studied later. Let us imagine an assemblage of N identical atoms arrayed on a fictitious crystal lattice with lattice constant a which is to be varied gradually from ∞ to the value a_0 of the real crystal. The electronic state $\phi_\mu(\mathbf{r} - \mathbf{R}_n)$ of an isolated atom at \mathbf{R}_n will resonate with that on the neighboring atoms $\mathbf{R}_{n'}$ and, due to the periodic array of atoms, better eigenstates should be running wave type linear combinations of them:

$$\psi_{\mu\mathbf{k}}(\mathbf{r}) = N^{-1/2} \sum_n \exp(i\mathbf{k} \cdot \mathbf{R}_n) \phi_\mu(\mathbf{r} - \mathbf{R}_n). \quad (7.1.1)$$

The corresponding eigenenergy $\varepsilon_\mu(\mathbf{k})$ is a continuous function of \mathbf{k} within the first Brillouin zone, forming an *energy band* μ similarly to the lattice vibrations we learned of in Chapter 5. This is the *tight binding approximation*, an approach to the *Bloch states* $\psi_{\mu\mathbf{k}}$ from the atomic states. Another approach, the *nearly free electron approximation*, starts from the free-electron states $\exp(i\mathbf{k} \cdot \mathbf{r})$ in a vacuum and takes account of the periodic potential of the crystal as a perturbation. We omit here the description of this method, referring the reader to standard text books on solids.^{1,2} There are available a number of more realistic methods of calculating the Bloch function $\psi_{\mu\mathbf{k}}(\mathbf{r})$, but, irrespective of the method, it can be written as a product of the form

$$\psi_{\mu\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}) u_{\mathbf{k}}(\mathbf{r}), \quad (7.1.2)$$

where $u_{\mathbf{k}}(\mathbf{r})$ is a periodic function such that $u_{\mathbf{k}}(\mathbf{r} + \mathbf{R}_n) = u_{\mathbf{k}}(\mathbf{r})$ for any lattice vector \mathbf{R}_n (see (5.1.15)). The validity of this *Bloch theorem* can be readily confirmed for the two typical approaches mentioned above. For instance, in our tight binding

approximation we have only to multiply (7.1.1) by $\exp(-i\mathbf{k} \cdot \mathbf{r})$, to obtain:

$$\exp(-i\mathbf{k} \cdot \mathbf{r})\psi_{\mu\mathbf{k}}(\mathbf{r}) = N^{-1/2} \sum_n \exp[-i\mathbf{k} \cdot (\mathbf{r} - \mathbf{R}_n)]\phi_{\mu}(\mathbf{r} - \mathbf{R}_n).$$

The r.h.s. is readily confirmed to be a periodic function of \mathbf{r} as $u_{\mathbf{k}}(\mathbf{r})$ above is, since replacing \mathbf{r} by $\mathbf{r} + \mathbf{R}_m$ amounts to renumbering of the lattice points from \mathbf{R}_n to $\mathbf{R}_n - \mathbf{R}_m$ which keeps the function invariant under the cyclic boundary condition. (See Refs. [1, 2] for a general proof.) Being energy eigenstates of an electron under a common periodic potential, the *normalized* Bloch functions for all (μ, \mathbf{k}) form an *orthonormal complete* set

$$(\psi_{\mu\mathbf{k}}, \psi_{\mu'\mathbf{k}'}) = \delta_{\mu,\mu'}\delta_{\mathbf{k},\mathbf{k}'}. \quad (7.1.3)$$

With a given set of Bloch functions, one can define *Wannier functions* by

$$a_{\mu}(\mathbf{r} - \mathbf{R}_n) \equiv N^{-1/2} \sum_n \exp(-i\mathbf{k} \cdot \mathbf{R}_n)\psi_{\mu\mathbf{k}}(\mathbf{r}). \quad (7.1.4)$$

In the particular case of the tight binding approximation (7.1.1) *without overlap* of atomic wave functions, $a_{\mu}(\mathbf{r} - \mathbf{R}_n)$ is nothing other than the atomic function $\phi_{\mu}(\mathbf{r} - \mathbf{R}_n)$, but are different for differently obtained $\psi_{\mu\mathbf{k}}(\mathbf{r})$ although the r.h.s. of (7.1.4) is always a function of $(\mathbf{r} - \mathbf{R}_n)$ due to (7.1.2). Since $\psi_{\mu\mathbf{k}}(\mathbf{r})$ multiplied by an arbitrary phase factor $\exp[i\alpha(\mathbf{k})]$ is also an eigenfunction, one can obtain different Wannier functions for different phase function $\alpha(\mathbf{k})$. Making use of this fact, one can minimize the spacial extension of the Wannier functions so as not to be too much extended compared with the atomic wave function. Let us assume this and consider them as *quasi*-atomic states. Wannier functions are convenient set of basis function since they satisfy the orthonormality relations

$$\int d\mathbf{r} a_{\mu}^*(\mathbf{r} - \mathbf{R}_n)a_{\mu'}(\mathbf{r} - \mathbf{R}_{n'}) = \delta_{\mu\mu'}\delta_{nn'}, \quad (7.1.5)$$

which are not satisfied by the overlapping atomic functions with $n \neq n'$.

For g -fold degenerate atomic states, the corresponding energy bands consist of the g branches with particular lines or points of degeneracy in the \mathbf{k} -space. The correspondence of the band to the atomic state from which it originates is obvious for inner atomic states, but not so clear for outermost states with bandwidths comparable to energy separations of the corresponding atomic states. A typical example is the covalent crystal of fourth-column elements such as diamond, silicon and germanium. The bonding and antibonding states between neighboring atoms which consist of their tetrahedral s-p hybridized orbitals extended towards each other constitute the occupied valence band and the empty conduction band, respectively. Even in this situation one can define the Wannier functions for the valence and conduction bands separately in terms of the corresponding Bloch states, not in terms of atomic states.

7.2 Band model for the classification of solids

The electrons are populated in the lowest possible energy bands at low temperatures. The electrical and optical properties of the crystal are completely different depending upon the occupancy of the highest populated band. If a finite bandgap ε_g exists between the highest filled-up band (valence band) μ and the lowest unoccupied band (conduction band) ν , as in molecular or ionic crystals with closed-shell structures of the constituents, or in the covalent crystals mentioned above, the electric current conveyed by the electron \mathbf{k} with group velocity $\partial\varepsilon_\mu(\mathbf{k})/\partial(\hbar\mathbf{k})$ is always canceled out by that of the electron $-\mathbf{k}$ with exactly opposite velocity up to that band μ , there being no electrons to be accelerated by the applied electric field unless an excitation energy at least equal to ε_g is supplied. As a result, the crystal behaves as an *insulator*.

However, if the bandgap is not large enough, as in silicon, germanium and many III–V compounds, some electrons are thermally excited from the valence band μ to the conduction band ν except at low temperatures, and the *carriers* of both signs, the *electrons* in the conduction band and the *positive holes* in the valence band, are readily accelerated by the applied electric field. As a result, the crystal behaves as an *intrinsic semiconductor*.

If the highest populated band is partly filled, as in alkali metals with outermost atomic s-shell half filled, the electrons at the Fermi energy which is well inside the band can be excited to empty states with different \mathbf{k} s by an infinitesimally small energy supply. Hence an appreciable fraction of the electrons contribute to the electric current under the applied electric field, typical of metallic conduction.

Alongside these, there are a variety of solids which are not appropriately described in terms of the one-electron band picture because of strong correlations among electrons, some of which will be mentioned in later chapters.

7.3 The roles of impurity atoms

Real crystals contain a significant number of impurity atoms and lattice defects. In insulators and semiconductors, an impurity atom with valency one greater than the host atom it replaces can readily supply the extra electron, which is no longer needed to participate in the host crystal bond, to the normally unoccupied band ν , contributing to the electric conduction. The impurity atom thus acts as a *donor*, and the crystal behaves as *extrinsic semiconductor* of n-type. As a result, the impurity atom is positively charged and can loosely capture the extra electron in its attractive Coulomb potential at a low enough temperature.³

In contrast, an impurity atom with valency one smaller than the host atom it replaces deprives one electron from the filled band μ (also called the *valence band*)

in order to complete the bond in the host crystal, thus acting as an *acceptor*, which amounts to supplying a positive hole to the otherwise filled band μ . The crystal behaves as an extrinsic semiconductor of p-type. The liberated positive hole can be loosely captured around the negatively charged impurity atom at a low enough temperature.

Intrinsic semiconductors become extrinsic semiconductors at low temperatures where the carriers excited across the gap are predominantly those supplied by the impurities.

7.4 Effective mass approximation for shallow impurity states

Shallow donor (acceptor) states mentioned in the previous section can be described⁴ as follows. The states have split off from the conduction band due to the attractive Coulomb potential $v(\mathbf{r}) = -e^2/4\pi\epsilon_c r$. Otherwise, there are no eigenstates of an electron within the bandgap (which is also called a *forbidden band* for this reason). In expanding the donor-state wave function on the basis of the complete orthogonal set of the Bloch functions of all bands, it is reasonable to suppose that the conduction band v makes the predominant contribution since energetically it is by far the band closest to the donor state. The Schrödinger equation for the donor-state wave function can then be written as

$$[h + v(\mathbf{r}) - E] \sum_{\mathbf{k}} f(\mathbf{k}) \psi_{v\mathbf{k}}(\mathbf{r}) = 0 \quad (7.4.1)$$

where h , the one-electron Hamiltonian given by $h = -(\hbar^2/2m)\nabla^2 +$ (periodic potential of the crystal due to all nuclei and all electrons in the filled bands), has the eigenenergies and eigenfunctions mentioned before:

$$[h - \varepsilon_{v\mathbf{k}}]\psi_{v\mathbf{k}}(\mathbf{r}) = 0. \quad (7.4.2)$$

It may be more intuitive to rewrite the wave function in (7.4.1) as a linear combination, with coefficients $F(\mathbf{R}_n)$, of the Wannier functions with the use of (7.1.4). Multiplying (7.4.1) by $a_v^*(\mathbf{r} - \mathbf{R}_n)$ and integrating on \mathbf{r} , one obtains the equations

$$\sum_n [h_{vnn'} + v_{nn'} - E\delta_{nn'}]F(\mathbf{R}_n) = 0 \quad (7.4.3)$$

where

$$h_{vnn'} \equiv \int d\mathbf{r} a_v^*(\mathbf{r} - \mathbf{R}_n) h a_v(\mathbf{r} - \mathbf{R}_{n'}) \equiv h_v(\mathbf{R}_n - \mathbf{R}_{n'}) \quad (7.4.4)$$

$$= N^{-1} \sum_{\mathbf{k}} \exp[i\mathbf{k} \cdot (\mathbf{R}_n - \mathbf{R}_{n'})] \varepsilon_v(\mathbf{k}), \quad (7.4.4a)$$

$$v_{nn'} \equiv \int d\mathbf{r} a_v^*(\mathbf{r} - \mathbf{R}_n) v(\mathbf{r}) a_v(\mathbf{r} - \mathbf{R}_{n'}) \cong \delta_{nn'} v_{nn}, \quad (7.4.5)$$

$$F(\mathbf{R}_n) \equiv N^{-1/2} \sum_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{R}_n) f(\mathbf{k}), \quad (7.4.6)$$

$$\sum_n |F(\mathbf{R}_n)|^2 = \sum_{\mathbf{k}} |f(\mathbf{k})|^2 = 1. \quad (7.4.7)$$

The first term of (7.4.3) can be rewritten as

$$\begin{aligned} \sum_{n'} h_v(\mathbf{R}_n - \mathbf{R}_{n'}) F(\mathbf{R}_{n'}) &= \sum_m h_v(\mathbf{R}_m) F(\mathbf{R}_n - \mathbf{R}_m) \\ &= \sum_m h_v(\mathbf{R}_m) \sum_p (p!)^{-1} [-\mathbf{R}_m \cdot \nabla_n]^p F(\mathbf{R}_n) \\ &= \sum_m h_v(\mathbf{R}_m) \exp[-\mathbf{R}_m \cdot \nabla_n] F(\mathbf{R}_n) \\ &= \varepsilon_v(-i\nabla_n) F(\mathbf{R}_n) \end{aligned} \quad (7.4.8)$$

where use has been made of the Taylor expansion of $F(\mathbf{R}_n - \mathbf{R}_m)$ at \mathbf{R}_n with ∇_n being the gradient with respect to \mathbf{R}_n , and the inverse transform of (7.4.4a):

$$\varepsilon_v(\mathbf{k}) = \sum_m h_v(\mathbf{R}_m) \exp[-i\mathbf{k} \cdot \mathbf{R}_m] \quad (7.4.9)$$

with \mathbf{k} being replaced by the operator $-i\nabla_n$.

Expand the energy of the conduction band around its bottom \mathbf{k}_c as

$$\varepsilon_v(\mathbf{k}) = \varepsilon_c + (\hbar^2/2)(\mathbf{k} - \mathbf{k}_c) \overline{\overline{m}}_c^{-1} (\mathbf{k} - \mathbf{k}_c) + \dots \quad (7.4.10)$$

where $\overline{\overline{m}}_c$ is the effective mass tensor. If the binding energy of the donor state is small, $F(\mathbf{R}_n)$ is a spatially extended slowly varying function. Hence, the difference equation (7.4.3) can be approximated by a differential equation with $\sum_{n'} h_{vnn'} F(\mathbf{R}_{n'})$ being replaced by $\varepsilon_v(-i\nabla) F(\mathbf{R})$ and, moreover, it is sufficient to consider, in the expansion (7.4.10), up to the quadratic term. After replacement $F(\mathbf{R}) = \exp(i\mathbf{k}_c \cdot \mathbf{R}) F'(\mathbf{R})$, one can write

$$\varepsilon_v(-i\nabla) F(\mathbf{R}) = \exp(i\mathbf{k}_c \cdot \mathbf{R}) [\varepsilon_c - (\hbar^2/2) \nabla (\overline{\overline{m}}_c)^{-1} \nabla] F'(\mathbf{R}).$$

Additionally, if $\mathbf{k}_c = 0$ and the effective mass is isotropic, the difference equation turns out to be of the same form as that for the hydrogen atom

$$[-(\hbar^2/2m_c) \nabla^2 - e^2/4\pi\epsilon_e R] F(\mathbf{R}) = (E - \varepsilon_c) F(\mathbf{R}), \quad (7.4.11)$$

except for the replacements of ϵ_0 (vacuum dielectric constant) by ϵ_e (dielectric constant contributed from bound electrons of the host crystal) and of m (electron mass) by m_c (the effective mass of the conduction band). The binding energy B_n and the orbital radius a_n of the donor-electron state with principal quantum number n is given by

$$B_n \sim Ry_d/n^2 (n = 1, 2, \dots), \quad Ry_d = (m_c/m)(\epsilon_0/\epsilon_e)^2 Ry, \quad (7.4.12)$$

$$a_n \sim a_d n^2, \quad a_d = (m/m_c)(\epsilon_e/\epsilon_0) a_H, \quad (7.4.13)$$

$$Ry = me^4/32\pi^2\epsilon_0^2\hbar^2 = 13.6 \text{ eV}, \quad a_H = 4\pi\epsilon_0\hbar^2/e^2m = 0.053 \text{ nm} \quad (7.4.14)$$

where Ry is the Rydberg constant expressed in units of energy and a_H is the orbital radius of the hydrogen atom.

In typical semiconductors such as Si, Ge and GaAs, the relative dielectric constant ϵ_c/ϵ_0 is of the order of 10 and the relative effective mass m_c/m of the conduction band is smaller than unity, so that the binding energy and the orbital radius in the ground state are of the order of 0.01 eV ($\ll \epsilon_g \sim 1$ eV) and 10 nm (\gg lattice constant ~ 0.3 nm), respectively. This justifies the use of the effective mass approximation mentioned above.

To elaborate, the bottom of the conduction band k_c is not at $k = 0$ in Si and Ge, being on [001] and equivalent directions (six valleys) in Si, and on the Brillouin zone edge point along [111] and equivalent directions (four valleys) in Ge, with longitudinal mass m_ℓ larger than the doubly-degenerate transverse mass m_t as is observed by cyclotron resonance experiments. While the equi-energy surface in the k -space of each valley is thus an *oblong* spheroid, the corresponding donor-state wave function in the \mathbf{r} -space is an *oblate* spheroid. The formula (7.4.12) for the hydrogen-like series is no longer valid. The correct wave functions of the donor ground state are linear combinations of those from the six (or four) valleys, the degeneracy being removed as the central cell correction is taken into account beyond the effective mass approximation.

7.5 Isoelectronic impurity

If the impurity atom has the same valency as the host atom it replaces, the bonds with the neighboring host atoms are already completed in its neutral state, no electron being liberated into the conduction band. However, if the electron affinity of the impurity atom, A_i , is sufficiently larger than that of the host atom, A_h , the impurity atom can capture another electron from the conduction band. Roughly speaking the difference $-(A_i - A_h)$ acts as a short-range attractive potential at the site of the impurity atom only. The problem is mathematically equivalent to the localized mode of lattice vibrations due to the presence of a low-mass impurity atom as described in Section 5.4, except that ω^2 in the latter corresponds to the *decreasing* direction of energy in the former (the localized mode appears at frequencies *above* the continuum of the host lattice phonons while the donor-bound state is at energies *below* the continuum of the conduction band). The bound state of the extra electron exists only when $(A_i - A_h)$ is comparable to the bandwidth. As will be mentioned later, an excitation, a bound electron-hole pair, is more readily bound by isoelectronic impurities.

8

Excitons

8.1 Frenkel excitons

When we considered low-energy elementary excitations such as optical phonons (Section 5.3) and the low-energy binding of donor electrons (Section 7.4), we described the background polarization effect of the electrons tightly bound to host atoms *phenomenologically* in terms of a constant susceptibility $\epsilon_e - \epsilon_0$ assuming that the polarizations have characteristic frequencies high enough to follow instantaneously the aforementioned low-frequency motions. In this chapter we will be concerned with the high-energy elementary excitations of tightly bound electrons contributing to susceptibility $\epsilon_e - \epsilon_0$ of the background polarization, in particular with excitons,¹ which is a first step to incorporating configuration interactions among one-electron excitations.

As the simplest model, consider an isolated atom (or ion) with one electron in the $1s$ state with energy ε_{1s} . An applied electric field gives rise to electronic polarization of the atom which quantum mechanically is due to the mixing of wave functions of odd parity np ($n = 2, 3, \dots$) states with energies ε_{np} to that of the even-parity $1s$ state. After the field is removed, the mixing coefficients contributing to the polarization oscillate with angular frequencies $\omega_n = (\varepsilon_{np} - \varepsilon_{1s})/\hbar$. Namely, the elementary excitations responsible for the polarization of an atom are excitations of its electron to higher states with different parity.

If identical atoms (or molecules) form an assemblage, the excitation of an atom resonates with those of neighboring atoms and, via the interatomic interactions, the excitation energy will propagate from atom to atom like a wave if the atoms are arrayed on a crystal lattice. The propagating wave of excitations in a molecular crystal was first conceived by Frenkel,² and the quantum of this elementary excitation is now called a *Frenkel exciton*.

One can describe this as follows. The ground-state wave function of the assemblage of identical atoms is given by

$$\Psi^{(\text{g})}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \prod_m \phi_s(\mathbf{r}_m - \mathbf{R}_m) \quad (8.1.1)$$

if one neglects wave function overlap among different atoms and ignores antisymmetrization among electronic coordinates. Within the same approximation, the wave function of the state in which only the n th molecule is in the excited state p is given by

$$\Phi_n(\mathbf{r}_1, \dots, \mathbf{r}_N) = \phi_p(\mathbf{r}_n - \mathbf{R}_n) \prod_{m(\neq n)} \phi_s(\mathbf{r}_m - \mathbf{R}_m). \quad (8.1.2)$$

The Hamiltonian of the system consists of the atomic Hamiltonians h_n and their interactions v_{nm} :

$$H = \sum_n h_n + \sum \sum_{n < m} v_{nm}. \quad (8.1.3)$$

If the atoms are arrayed on a crystal lattice, the expectation value of the energy in the excited state, $H_{nn}^{(\text{e})} \equiv (\Phi_n, H \Phi_n)$, should be independent of n due to the translational symmetry. Its difference ε from the ground-state energy $E^{(\text{g})}$ of the crystal is given by the energy difference $(\varepsilon_p - \varepsilon_s)$ between the excited and ground states of an isolated atom corrected by the difference $(\Delta_p - \Delta_s)$ of their environmental energies due to the vs .

However, (8.1.3) has matrix elements between different Φ_n s due to the interatomic interactions, namely

$$\begin{aligned} H_{nm}^{(\text{e})} \equiv (\Phi_n, H \Phi_m) &= \int d\mathbf{r}_n \int d\mathbf{r}_m \phi_p(\mathbf{r}_n - \mathbf{R}_n) \phi_s(\mathbf{r}_m - \mathbf{R}_m) v_{nm} \\ &\times \phi_s(\mathbf{r}_n - \mathbf{R}_n) \phi_p(\mathbf{r}_m - \mathbf{R}_m). \end{aligned} \quad (8.1.4)$$

All other terms vanish due to the orthogonality of ϕ_s and ϕ_p . In v_{nm} , consisting of the interactions between an electron and a nucleus of the n th atom at \mathbf{R}_n and those of the m th atom at \mathbf{R}_m , only the electron–electron Coulomb interaction: $e^2/4\pi\epsilon_0|\mathbf{r}_n - \mathbf{r}_m|$ remains to contribute to the matrix element (8.1.4) since the electron–nucleus and internuclear interactions contribute nothing because of the orthogonality of ϕ_s and ϕ_p . Assuming the spacial extensions of the wave functions ϕ_s and ϕ_p to be small compared to the interatomic distance $|\mathbf{R}_{nm}| \equiv |\mathbf{R}_n - \mathbf{R}_m|$, one can make multipole expansions of the product $\phi_p \phi_s$ around the respective centers in the integral (8.1.4). The first term is the *dipole–dipole* interaction, which is the only one we will consider here. Denoting the transition matrix element of the electric dipole by

$$\mu \equiv \int d\mathbf{r} \phi_p(\mathbf{r})(-e\mathbf{r})\phi_s(\mathbf{r}), \quad (8.1.5)$$

one can write (8.1.4) as the dipole–dipole interaction at distance \mathbf{R}_{nm} :

$$H_{nm} = [\mu^2/R_{nm}^3 - 3(\boldsymbol{\mu} \cdot \mathbf{R}_{nm})^2/R_{nm}^5]/4\pi\epsilon_0 \equiv D(\mathbf{R}_{nm}). \quad (8.1.6)$$

As in the case of lattice vibrations, the Hamiltonian $H^{(e)}$ can be diagonalized by taking the running wave-like linear combination of Φ_n :

$$\Psi_{\mathbf{K}}^{(e)} = N^{-1/2} \sum_n \exp(i\mathbf{K} \cdot \mathbf{R}_n) \Phi_n \quad (8.1.7)$$

with diagonalized energy

$$E_{\mathbf{K}}^{(e)} = (\Psi_{\mathbf{K}}^{(e)}, H \Psi_{\mathbf{K}}^{(e)}) = E^{(g)} + (\varepsilon + D_{\mathbf{K}}), \quad (8.1.8)$$

$$D_{\mathbf{K}} \equiv \sum_{n(\neq 0)} D(\mathbf{R}_n) \exp(-i\mathbf{K} \cdot \mathbf{R}_n). \quad (8.1.9)$$

Equation (8.1.7) indicates that eigenstates in the *subspace of one-electron excitation* are the waves in which excitation propagates from atom to atom with definite wave vectors \mathbf{K} . What *propagates* here is *not the electron but the energy of excitation* within the present approximation which neglects the overlap of wave functions at different atoms (the effect of overlap will be considered in Sections 8.2 and 8.3). The excitation energy $E_{\mathbf{K}} = \varepsilon + D_{\mathbf{K}}$ of one-electron excitation is a continuous function of the wave vector \mathbf{K} , forming the *exciton band*. Its energy width is due to the interactions between the transition dipole moments of different atoms.

Dipolar interaction $D(\mathbf{R}_{nm})$ is a *long-range* interaction which decays as R_{nm}^{-3} , and the summation (8.1.9) on the three-dimensional lattice points sometimes causes a singular convergence depending on the direction $\mathbb{K} \equiv \mathbf{K}/K$ in the vanishing limit of K .³ Let us take K to be much smaller than the reciprocal lattice b , and consider a *macroscopically small* sphere of radius R_c such that $a \ll R_c \ll K^{-1}$ around the origin $\mathbf{R}_n = 0$ where a is the lattice constant (see eq. (5.1.17) for as and bs and their relations). Divide the summation on n to the inner region ($R_n < R_c$) and the outer region ($R_n > R_c$). One can replace the outer-region summation by integration: $N_0 \int d\mathbf{R} \cdots$ (where $N_0 \sim a^{-3}$ is the number of atoms per unit volume) since $Ka \ll 1$. With the use of polar coordinates with \mathbf{K} along the z -axis and with $\boldsymbol{\mu}$ to be on the z - x plane from which the longitude angle ϕ is to be measured, the integral is given by

$$\begin{aligned} I &= (N_0 \mu^2 / 4\pi \epsilon_0) \int_{R_c}^{\infty} dR/R \int_0^{\pi} \exp(-iK R \cos \theta) \sin \theta d\theta \\ &\quad \times \int_{-\pi}^{+\pi} (1 - 3 \cos^2 \theta') d\phi \\ &= -(N_0 \mu^2 / 3\epsilon_0) [(1 - 3 \cos^2 \theta_0) + O[(K R_c)^2]] \end{aligned}$$

where θ' is the angle between $\boldsymbol{\mu}(\theta_0, 0)$ and $\mathbf{R}(\theta, \phi)$ given by $\cos \theta' = \cos \theta_0 \cos \theta + \sin \theta_0 \sin \theta \cos \phi$. Since one can put $\exp(-i\mathbf{K} \cdot \mathbf{R}_n) \cong 1$ in the inner region, one finally obtains

$$D_{\mathbf{K}} = \sum_{R_n < R_c} D(\mathbf{R}_n) - (N_0/3\epsilon_0)[\mu^2 - 3(\boldsymbol{\mu} \cdot \mathbb{K})^2] \quad (Ka \ll 1). \quad (8.1.10)$$

The independence of integral I on the choice of lower bound R_c assures the independence of summation in (8.1.10) on the upper bound R_c . It should be mentioned in passing that this summation vanishes for any lattice with *cubic* symmetry. Note that (8.1.10) depends on the direction \mathbb{K} , the singular convergence as we mentioned before.

However, we must mention also the less singular situations which occur rather often. As a first example, let us note that the atomic p-states of our model are in fact triply degenerate, and we must generalize (8.1.10) to 3×3 matrices corresponding to p_x , p_y and p_z components, each with equal length $\mu = |\boldsymbol{\mu}|$ of the transition dipole moment from the s-like ground state. Then the ij-component of this matrix for any lattice with cubic symmetry is given by

$$(D_{\mathbf{K}})_{ij} = -(N_0\mu^2/3\epsilon_0)[\delta_{ij} - 3K_i K_j / K^2], \quad (8.1.11)$$

which is diagonalized by taking the coordinate system with the z -axis along \mathbb{K} . In fact, one obtains one longitudinal wave with $\boldsymbol{\mu} \parallel \mathbb{K}$ and energy $(D_{\mathbf{K}})_{33} = +(2/3)(N_0\mu^2/\epsilon_0)$, and two transverse waves with $\boldsymbol{\mu} \perp \mathbb{K}$ and energy $(D_{\mathbf{K}})_{11} = (D_{\mathbf{K}})_{22} = -(1/3)(N_0\mu^2/\epsilon_0)$. Their energy difference, $N_0\mu^2/\epsilon_0$, exactly corresponds to the force constant difference $N_0e'^2/\epsilon_e$ between longitudinal and transverse optical modes (see (5.3.15)), both originating from the depolarizing field. The convergence singularity is removed by this ℓ - t splitting since the energy or force constant within each of the ℓ - and t -branches is now a continuous function of \mathbf{K} with no anomaly at $\mathbf{K} = 0$.

For the second example, we should mention that many organic molecular crystals consist of unit cells containing *more than two identical molecules with different orientations*, presumably because the molecules with structural symmetry lower than an atom can form an energetically favorable close-packed crystal structure by so doing. Even when the excited state of a molecule is non-degenerate, the dipolar interaction $D_{\mathbf{K}}$ is a matrix with the dimensionality being the number of molecules in a unit cell. This results in the so-called *Davydov splitting* of the absorption spectra,⁴ as has been confirmed in a number of aromatic molecular crystals. The convergence singularity in the neighborhood of $\mathbf{K} = 0$ is removed in many of the cases of Davydov splitting studied so far.

8.2 Wannier excitons

We are well aware of the fact that atoms and molecules have a host of higher excited states tending to the Rydberg states in which the excited electron is still bound in the Coulomb potential arising from the ion left behind. At still higher energies they exhibit the ionization continuum.

In the case of a crystalline assemblage of atoms or molecules, one-electron states are the Bloch states extended throughout the entire space of the crystal and forming continuous energy bands. In insulating crystals characterized by the existence of a finite *bandgap* ε_g between the top of the valence band and the bottom of the conduction band, the ionization continuum of an atom or a molecule corresponds to excitation of an electron from the valence band to the conduction band or, equivalently, to creation of a free pair of an electron in the conduction band and a hole in the valence band, while the Rydberg states correspond to Coulomb-bound states of an electron-hole pair with energy given by $\varepsilon_g - B_n$. The binding energy B_n is expected to tend to those of the Rydberg-like series, and the orbital radius a_n of the electron-hole relative motion to the hydrogen-like ones, with the difference that the electronic mass m is to be replaced by the reduced mass $m_r = (m_e^{-1} + m_h^{-1})^{-1}$ of the effective masses m_e and m_h of an electron in the conduction band and a hole in the valence band, and the dielectric constant ϵ_0 is to be replaced by ϵ_e of the crystal:

$$B_n \sim Ry_{\text{ex}}/n^2 (n = 1, 2, \dots), \quad Ry_{\text{ex}} = (m_r/m)(\epsilon_0/\epsilon_e)^2 Ry, \quad (8.2.1)$$

$$a_n \sim a_{\text{ex}} n^2, \quad a_{\text{ex}} = (m/m_r)\epsilon_e/\epsilon_0 a_H, \quad (8.2.2)$$

analogous to eqs. (7.4.12) to (7.4.14) for the donor electron.

The bound electron-hole pair is called a *Wannier exciton*. The hydrogen-like series of the Wannier exciton has been observed as line spectra in several insulators (alkali halides, cuprous halides, rare-gas solids, Cu_2O , etc.) whose conduction-band bottom and valence-band top are at the same point, $\mathbf{k} = 0$, of the Brillouin zone, each with isotropic effective masses of m_e and m_h (except for some complications due to degeneracy of the valence band), though such an idealized situation is not realized in other insulators. Since ϵ_e/ϵ_0 is as large as ten in insulators with not too large ε_g , the binding energy of even a 1s exciton is as small as 10^{-1} – 10^{-2} eV.

An assumption underlying (8.2.1, 8.2.2) that the electron-hole pair is subject to the screened Coulomb potential: $-e^2/4\pi\epsilon_e|\mathbf{r}_e - \mathbf{r}_h|$ as if it were in the *dielectric continuum* is justified only if the exciton radius a_n is much larger than the lattice constant (see Section 7.4). A significant deviation from the hydrogen-like series is found for the lowest lines (small n) even in these idealistic crystals. The Frankel exciton described in Section 8.1 can be considered as an extreme example of such a deviation.

As another criterion, the use of ϵ_e is justified since the binding energy B_n is much smaller than ϵ_g which is the smallest binding energy of host bound electrons in an insulator (see Section 9.3). For exciton binding energy even smaller than the energy of a longitudinal optical phonon, the polaron effects of the individual electron and hole must be taken into account, with ϵ_e to be replaced by the static dielectric constant $\epsilon(0)(> \epsilon_e)$, as will be mentioned in Section 9.2.

We have been concerned thus far only with bound states of an electron–hole pair. The pair also has ionized states with continuous energy characterized by the wave vector \mathbf{k} of the relative motion. The pair can also make a translational motion with mass $m_{\text{ex}} = m_e + m_h$ and wave vector \mathbf{K} , irrespectively of bound or ionized states. Therefore, the total energy of an electron–hole pair consists of energy E'_λ for the relative motion with quantum number $\lambda = (n, l, m)$ or \mathbf{k} and translational kinetic energy with quantum number \mathbf{K} :

$$E(\lambda, \mathbf{K}) = \epsilon_g + E'_\lambda + \hbar^2 \mathbf{K}^2 / 2m_{\text{ex}}, \quad (8.2.3)$$

$$E'_\lambda = -B_n \text{ (bound states), } = +\hbar^2 \mathbf{k}^2 / 2m_r \text{ (ionized states).} \quad (8.2.4)$$

8.3 Elementary excitations in many-electron system of insulator

In the preceding sections, we have described two aspects of an exciton, one as a propagating wave of intra-atomic (-molecular) excitation energy, and another as a bound pair of a conduction-band electron and a valence-band hole. The two aspects can be incorporated into a general theory in which the exciton is considered as an elementary excitation in the many-electron system of an insulator.⁵ In the ground state Φ_g of an insulator, all lower bands up to the band μ are completely occupied by electrons while all upper bands starting from the band ν are empty. In the Hartree–Fock approximation, Φ_g is expressed as a *Slater determinant* of all Bloch functions of the lower bands (so as to change sign against the exchange of coordinates of any two electrons in accordance with the Pauli exclusion principle). According to the band picture, the subspace of one-electron excitation is spanned by all the states $\Psi(\nu, \mathbf{k}_1 \leftarrow \mu, \mathbf{k}_2)$ (where ν : conduction band, μ : valence band, both being assumed to be degenerate; $\mathbf{k}_1, \mathbf{k}_2$ in the first Brillouin zone) each of which is obtained by replacing the Bloch function $\psi_{\mu\mathbf{k}_2}$ by $\psi_{\nu\mathbf{k}_1}$ in the above-mentioned Slater determinant.

The energy eigenstate should be specified by total wave vector $\mathbf{K} = \mathbf{k}_1 - \mathbf{k}_2$ due to the translational symmetry of the system (the minus (–) sign before \mathbf{k}_2 is due to the fact that the electron with wave vector \mathbf{k}_2 has been *removed*). Hence, we write $\mathbf{k}_1 = \mathbf{k}$ and $\mathbf{k}_2 = \mathbf{k} - \mathbf{K}$ and consider the set of all \mathbf{k} for a fixed value of \mathbf{K} . Let us note that the state $\Psi(\nu\mathbf{k} \leftarrow \mu\mathbf{k} - \mathbf{K})$ has excitation energy $\epsilon_{\nu\nu}(\mathbf{k}) - \epsilon_{\mu\mu}(\mathbf{k} - \mathbf{K})$ in the one-electron picture in which the interaction between the electron and the

positive hole left behind is neglected. Here $\varepsilon_{vv'}(\mathbf{k})$ is the energy matrix for the degenerate conduction band which is a generalization of $\varepsilon_v(\mathbf{k})$ for the non-degenerate conduction band v .

Due to the electron-hole interaction, however, one has to consider the *configuration interaction*. Namely, the energy eigenstates of the many-electron system within the subspace of one-electron excitation should be a *linear combination* of $\Psi_{v\mu\mathbf{k}\mathbf{K}} \equiv \Psi(v\mathbf{k} \leftarrow \mu\mathbf{k} - \mathbf{K})$ with respect to $(v\mu\mathbf{k})$. One can derive, on the basis of the many-electron theory within the Hartree-Fock approximation, the energy matrix $H(\mathbf{K})_{v\mu\mathbf{k},v'\mu'\mathbf{k}'}$ for the set of functions $\Psi_{v\mu\mathbf{k}\mathbf{K}}$ such that the eigenvalue $E_{\lambda\mathbf{K}}$ of $H(\mathbf{K})$ with eigenfunction $\Psi_{\lambda\mathbf{K}}$ gives the *excitation energy* (the energy measured from that of the ground state) of the interacting electron-hole pair:

$$\sum_{v'\mu'\mathbf{k}'} H(\mathbf{K})_{v\mu\mathbf{k},v'\mu'\mathbf{k}'} f_{\lambda\mathbf{K}}(v'\mu'\mathbf{k}') = E_{\lambda\mathbf{K}} f_{\lambda\mathbf{K}}(v\mu\mathbf{k}), \quad (8.3.1)$$

$$\Psi_{\lambda\mathbf{K}} = \sum_{v\mu\mathbf{k}} f_{\lambda\mathbf{K}}(v\mu\mathbf{k}) \Psi_{v\mu\mathbf{k}\mathbf{K}}. \quad (8.3.2)$$

Leaving a somewhat lengthy calculation to Appendix 1, we give here only the result:

$$H(\mathbf{K})_{v\mu\mathbf{k},v'\mu'\mathbf{k}'} = \delta_{\mathbf{k}\mathbf{k}'} [\delta_{\mu\mu'} \varepsilon_{vv'}(\mathbf{k}) - \delta_{vv'} \varepsilon_{\mu\mu'}(\mathbf{k} - \mathbf{K})] \\ - v_{v\mathbf{k}\mu'\mathbf{k}' - \mathbf{K}\mu\mathbf{k} - \mathbf{K}v'\mathbf{k}'} + v_{\mu'\mathbf{k}' - \mathbf{K}v\mathbf{k}\mu\mathbf{k} - \mathbf{K}v'\mathbf{k}'}, \quad (8.3.3)$$

$$v_{\lambda_1\mathbf{k}_1\lambda_2\mathbf{k}_2\lambda_3\mathbf{k}_3\lambda_4\mathbf{k}_4} \equiv \int \int d\mathbf{r} d\mathbf{r}' \psi_{\lambda_1\mathbf{k}_1}^*(\mathbf{r}) \psi_{\lambda_2\mathbf{k}_2}^*(\mathbf{r}') v(\mathbf{r} - \mathbf{r}') \\ \times \psi_{\lambda_3\mathbf{k}_3}(\mathbf{r}') \psi_{\lambda_4\mathbf{k}_4}(\mathbf{r}). \quad (8.3.4)$$

One notes that the terms $-v \dots$ and $+v \dots$ in (8.3.3) represent the Coulomb and exchange interactions, respectively, between the electron and the positive hole. Namely, the electron and the hole move in their respective bands with Coulomb *attraction* and exchange *repulsion*, their sign reversal being due to the positive hole instead of the second electron.

It may be more intelligible to transform the Bloch basis to the Wannier basis with the use of transformations (7.1.4) and (7.4.6). One can then write the eigenequation (8.3.1) as

$$\sum_{v'\mu'\mathbf{\ell}'} H(\mathbf{K})_{v\mu\mathbf{\ell}v'\mu'\mathbf{\ell}'} F_{\lambda\mathbf{K}}(v'\mu'\mathbf{\ell}') = E_{\lambda\mathbf{K}} F_{\lambda\mathbf{K}}(v\mu\mathbf{\ell}), \quad (8.3.5)$$

$$H(\mathbf{K})_{v\mu\mathbf{\ell}v'\mu'\mathbf{\ell}'} \equiv N^{-1} \sum_{\mathbf{k}\mathbf{k}'} \exp(i\mathbf{k} \cdot \mathbf{R}_{\mathbf{\ell}}) H(\mathbf{K})_{v\mu\mathbf{k},v'\mu'\mathbf{k}'} \exp(-i\mathbf{k}' \cdot \mathbf{R}_{\mathbf{\ell}'}) \\ = [\delta_{\mu\mu'} \varepsilon_{vv'}(-i\nabla_{\mathbf{\ell}}) - \delta_{vv'} \varepsilon_{\mu\mu'}(-i\nabla_{\mathbf{\ell}} - \mathbf{K})] \\ - \delta_{\mathbf{\ell}\mathbf{\ell}'} v_{vv'\mu\mu'}(\mathbf{R}_{\mathbf{\ell}}) + \delta_{\mathbf{\ell}\mathbf{0}} \delta_{\mathbf{\ell}'\mathbf{0}} \sum_m \exp(-i\mathbf{K} \cdot \mathbf{R}_m) v_{v\mu v'\mu'}(\mathbf{R}_m), \quad (8.3.6)$$

$$v_{vv'\mu\mu'}(\mathbf{R}_{\mathbf{\ell}}) \equiv \int \int d\mathbf{r} d\mathbf{r}' a_v^*(\mathbf{r} - \mathbf{R}_{\mathbf{\ell}}) a_{v'}(\mathbf{r} - \mathbf{R}_{\mathbf{\ell}}) v(\mathbf{r} - \mathbf{r}') a_{\mu}(\mathbf{r}') a_{\mu'}^*(\mathbf{r}'). \quad (8.3.7)$$

where ∇_ℓ denotes the gradient with respect to \mathbf{R}_ℓ . In deriving the interaction terms in (8.3.6) from those in (8.3.3), we have considered only those terms with $n_1 = n_4$ and $n_2 = n_3$ in the integral

$$\int \int d\mathbf{r} d\mathbf{r}' a_{\lambda_1}^*(\mathbf{r} - \mathbf{R}_{n_1}) a_{\lambda_2}^*(\mathbf{r}' - \mathbf{R}_{n_2}) v(\mathbf{r} - \mathbf{r}') a_{\lambda_3}(\mathbf{r}' - \mathbf{R}_{n_3}) a_{\lambda_4}(\mathbf{r} - \mathbf{R}_{n_4})$$

because of the small overlap of Wannier functions at different sites. Note that we have chosen the definition for the matrix elements of v in (8.3.7) to be different from that of (8.3.4) as regards the order of suffixes of the wave functions, for the sake of better insight into the physical meaning as seen below.

As is obvious by inspecting the band-diagonal elements ($v = v'$, $\mu = \mu'$), the first interaction term in (8.3.6) represents the attractive Coulomb interaction between the electron excited to the site \mathbf{R}_ℓ and the hole left behind. The second interaction term originating from the exchange term (the last term in (8.3.3)) represents the Fourier transform of the dipole–dipole interaction corresponding to (8.1.9), with its prefactor $\delta_{\ell 0} \delta_{\ell' 0}$ indicating that the dipole transition is confined within the same site due to the above approximation of neglecting the overlap of Wannier functions.

One can readily take into account the electron spin, so far ignored, by interpreting the above v and μ to include implicitly the spin variables, or more explicitly by replacing $v \rightarrow (v, \tau)$ and $\mu \rightarrow (\mu, \sigma)$ and interpreting the integration on the spatial coordinates \mathbf{r} to include the sum on the spin variables τ (or σ) which can take $\pm(1/2)$. In the absence of spin–orbit interaction (to be described in Section 8.5), the one-electron energy $\varepsilon_{v v'}$ is independent of the spin variable and one can write the r.h.s. of eq. (8.3.6) as

$$\begin{aligned} & \delta_{\tau\tau'} \delta_{\sigma\sigma'} [\delta_{\mu\mu'} \varepsilon_{v v'} (-i \nabla_\ell) - \delta_{v v'} \varepsilon_{\mu\mu'} (-i \nabla_\ell - \mathbf{K}) \\ & - \delta_{\ell\ell'} v_{v v' \mu \mu'}(\mathbf{R}_\ell)] + \delta_{\tau\sigma} \delta_{\tau'\sigma'} \delta_{\ell\ell'} \delta_{\ell' 0} W_{v \mu v' \mu'}(\mathbf{K}), \end{aligned} \quad (8.3.8)$$

where we have used the Fourier transform of the exchange integral:

$$W_{v \mu v' \mu'}(\mathbf{K}) \equiv \sum_m \exp(-i \mathbf{K} \cdot \mathbf{R}_m) v_{v \mu v' \mu'}(\mathbf{R}_m). \quad (8.3.9)$$

While the spin prefactor of the first term of (8.3.8), $\delta_{\tau\tau'} \delta_{\sigma\sigma'}$, is the “unit” or “identity” operator $\mathbb{1} = \mathbb{1}_e \times \mathbb{1}_h$ in the two-electron spin space, that of the second term, $\delta_{\tau\sigma} \delta_{\tau'\sigma'}$, has to be diagonalized in this space in order to obtain the energy eigenvalue of (8.3.8). This can be done in terms of the spin singlet and triplet states as follows. In the present space of one electron and one hole, the total spin is given by $S = \tau - \sigma$ since the hole has spin $-\sigma$ where $(\mu\sigma)$ is the state from which the electron has been removed. With the spin states $|\tau\sigma\rangle$ as basis, the triplet ($S = 1$) and singlet

($S = 0$) states are given, somewhat differently from the two-electron problem, by

$$\begin{aligned} S_Z = +1 : & \quad -|\uparrow\downarrow\rangle, \\ S = 1 \quad S_Z = 0 : & \quad (|\uparrow\uparrow\rangle - |\downarrow\downarrow\rangle)/\sqrt{2}, \end{aligned} \quad (8.3.10)$$

$$\begin{aligned} S_Z = -1 : & \quad |\downarrow\uparrow\rangle, \\ S = 0 & \quad (|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)/\sqrt{2}. \end{aligned} \quad (8.3.10a)$$

One can readily confirm, by taking matrix elements of $\delta_{\tau\sigma}\delta_{\tau'\sigma'}$ with respect to the basis functions of (8.3.11), that it is diagonalized as $2\delta_{S,0}$. Thus the Hamiltonian (8.3.8) is diagonalized in the spin space as

$$\begin{aligned} & [\delta_{\mu\mu'}\varepsilon_{vv'}(-i\nabla_\ell) - \delta_{vv'}\varepsilon_{\mu\mu'}(-i\nabla_\ell - \mathbf{K}) - \delta_{\ell\ell'}v_{vv'\mu\mu'}(\mathbf{R}_\ell)] \\ & + 2\delta_{S,0}\delta_{\ell 0}\delta_{\ell' 0}W_{v\mu v'\mu'}(\mathbf{K}). \end{aligned} \quad (8.3.11)$$

By far the most dominant among the terms in (8.3.9) is the band-diagonal element of the intra-atomic exchange term

$$v_{v\mu v\mu}(0) = \int \int d\mathbf{r} d\mathbf{r}' a_v^*(\mathbf{r}) a_\mu(\mathbf{r}) v(\mathbf{r} - \mathbf{r}') a_v(\mathbf{r}') a_\mu^*(\mathbf{r}') \quad (8.3.12)$$

(see (8.3.7)). It is positive definite because of its form of Coulomb interaction of the complex charge density $\rho(\mathbf{r}) \equiv a_v^*(\mathbf{r}) a_\mu(\mathbf{r})$ with its complex conjugate $\rho^*(\mathbf{r})$. Hence the singlet exciton is energetically always higher than the triplet exciton.

To be more exact, one must use the screened Coulomb potential: $v(\mathbf{r} - \mathbf{r}') = e^2/4\pi\epsilon_e|\mathbf{r} - \mathbf{r}'|$ in the *Coulomb* integral (8.3.7) for $\mathbf{R}_\ell \neq 0$ instead of the unscreened one: $v(\mathbf{r} - \mathbf{r}') = e^2/4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|$, since the angular frequency of the electron-hole relative motion in an exciton ($\sim R_{y_{\text{ex}}}/\hbar$) is much *slower* than that of the motion of bound electrons in the host crystal ($\sim \varepsilon_g/\hbar$). The same situation has been seen with the *slow* motion of atoms (lattice vibrations) and *shallow* bound donor electrons. This screening, which is introduced here phenomenologically, belongs to the many-body effect which is beyond the scope of the one-electron excitation model considered here. The theoretical foundation of this screening effect will be given in Section 9.3 in terms of the electronic polaron. The screening for the *exchange* integral is a more subtle problem and will be studied in two steps: later in this section and in Section 8.4.

Let us solve the electronic eigenvalue problem of the Hamiltonian (8.3.11), treating the second term as a small perturbation. For simplicity, we consider isotropic crystals and moreover confine ourselves to a model case where

$$\begin{aligned} \varepsilon_{vv'}(\mathbf{k}) &= \delta_{vv'}\varepsilon_c(\mathbf{k}), & \varepsilon_c(\mathbf{k}) &= \varepsilon_c + \hbar^2\mathbf{k}^2/2m_e, \\ \varepsilon_{\mu\mu'}(\mathbf{k}) &= \delta_{\mu\mu'}\varepsilon_v(\mathbf{k}), & \varepsilon_v(\mathbf{k}) &= \varepsilon_v - \hbar^2\mathbf{k}^2/2m_h, \end{aligned} \quad (8.3.13)$$

It is convenient to expand the one-electron energy in the Taylor series around its minimum $\mathbf{k}_m(\mathbf{K}) \equiv (m_e/M)\mathbf{K}$ for a given \mathbf{K} , as

$$\begin{aligned}\varepsilon_c(\mathbf{k}) - \varepsilon_v(\mathbf{k} - \mathbf{K}) &= \varepsilon^{(m)}(\mathbf{K}) + (\hbar^2/2)(\mathbf{k} - \mathbf{k}_m)m_r(\mathbf{K})^{-1}(\mathbf{k} - \mathbf{k}_m) + \dots \\ &= \varepsilon_c - \varepsilon_v + \hbar^2\mathbf{K}^2/2m_{ex} + \hbar^2(\mathbf{k} - \mathbf{k}_m)^2/2m_r,\end{aligned}\quad (8.3.14)$$

where translational mass m_{ex} and reduced mass m_r are defined by

$$m_{ex} \equiv m_e + m_h, \quad m_r^{-1} \equiv m_e^{-1} + m_h^{-1}. \quad (8.3.15)$$

Then, rewriting $F_{\lambda\mathbf{K}}(\mathbf{R}) = \exp(i\mathbf{k}_m(\mathbf{K}) \cdot \mathbf{R})F_{\lambda}^l(\mathbf{R})$, one can solve the eigenvalue problem of the unperturbed part of (8.3.11) as the effective mass equation for the relative motion:

$$[-(\hbar^2/2m_r)\nabla^2 - v(\mathbf{R})]F_{\lambda}^l(\mathbf{R}) = E_{\lambda 0}^l F_{\lambda}^l(\mathbf{R}). \quad (8.3.16)$$

Then the eigenenergy and eigenfunction of the two-body problem are given respectively by the sum of the energies

$$E_{\lambda\mathbf{K}} \equiv \hbar^2\mathbf{K}^2/2m_{ex} + E_{\lambda 0}^l, \quad (8.3.17)$$

and the product $F_{\lambda\mathbf{K}}(\mathbf{R})$ of the eigenfunctions of the translational and relative motions.

The contribution of the last term of (8.3.11), given by

$$2\delta_{s,0}|F_{\lambda}^l(0)|^2 W_{\nu\mu\nu'\mu'}(\mathbf{K}) \quad (8.3.18)$$

in the first-order perturbation theory, is not yet diagonalized with respect to the band indices ν and μ . Let us confine ourselves to the optically allowed transition $\nu \rightleftharpoons \mu$, the only case in which *interatomic* terms of $W_{\nu\mu\nu'\mu'}(\mathbf{K})$ are non-vanishing. Consider the $s \rightleftharpoons p$ transition as a typical example, and rewrite the suffix μ for p_x, p_y, p_z bands as $j = 1, 2, 3$ to avoid confusion with the transition dipole moment μ . Then the first term of (8.3.9), the *intra-atomic* exchange energy, is diagonal:

$$v_{sjsj'}(0) = \delta_{jj'} \int \int d\mathbf{r} d\mathbf{r}' a_s^*(\mathbf{r}) a_j(\mathbf{r}) v(\mathbf{r} - \mathbf{r}') a_s(\mathbf{r}') a_j^*(\mathbf{r}') \quad (8.3.19)$$

(see also (8.3.12)), while the sum of the *interatomic* terms, the dipole–dipole interactions, can be diagonalized as in to Section 8.1, to give

$$D_{\mathbf{K}t} = -(1/3)N_0\mu^2/\epsilon_{ce}, \quad D_{\mathbf{K}\ell} = +(2/3)N_0\mu^2/\epsilon_{ce}, \quad (8.3.20)$$

where t and ℓ denote the transverse ($\mu \perp \mathbf{K}$) and longitudinal ($\mu \parallel \mathbf{K}$) excitons, respectively. As the screening constant we have *provisionally* used the dielectric constant ϵ_{ce} contributed by atomic *core electrons* only, because we are considering the optical transitions of *valence electrons* in a background medium with dielectric

constant ϵ_{ce} . However, it has to be corrected by contributions of *other transitions* of the valence electrons, as will be seen in Section 8.4.

8.4 Optical spectra of excitons

We are now in a position to study the fundamental absorption spectra of an insulator due to the one-electron excitations described in the preceding sections. Since we have made use of the real Hamiltonian with long-range Coulomb interactions explicitly considered, we can resort to either of the \mathbf{E} (transverse case, (6.3.2)) and \mathbf{D} methods mentioned in Section 6.3 if we confine ourselves to isotropic crystals. Here we start with the \mathbf{E} method, which is directly connected with the optical spectra, and relate it to the \mathbf{D} method in a more elaborate study of longitudinal excitons in this section and in the study of the motion of a charged particle in the field of these longitudinal excitons in Section 9.3.

In the absence of spin-orbit interaction, one can remove the individual spin indices and confine ourselves to the singlet excited states $S = 0$ which alone contribute to the electric susceptibility. Denoting by \mathbf{q} the wave vector *in the crystal* of an incoming photon with frequencies ω , one can write the corresponding susceptibility (6.1.31) as

$$\begin{aligned}\chi_t(\mathbf{q}, \omega) &= \chi_t(\mathbf{q}, \omega)_{ii} \\ &= 2 \sum_{\lambda \mathbf{K}} [(P_{i,q})_{g,\lambda \mathbf{K}} (P_{i,-q})_{\lambda \mathbf{K},g} / (E_{0\lambda \mathbf{K}} - \hbar\omega - i\gamma) \\ &\quad + (P_{i,-q})_{g,\lambda \mathbf{K}} (P_{i,q})_{\lambda \mathbf{K},g} / (E_{0\lambda \mathbf{K}} + \hbar\omega + i\gamma)],\end{aligned}\quad (8.4.1)$$

where we have considered a single ground state and transverse waves of one-electron excited states $\lambda \mathbf{K}$, ignoring the effects of lattice vibrations which will be considered in Chapter 10. The factor 2 on the r.h.s. of (8.4.1) is for the contribution of two electrons with opposite spin, and $E_{0\lambda \mathbf{K}}$ denotes the energy of a singlet transverse exciton with relative motion quantum number λ and translational wave vector \mathbf{K} . Writing down the identity $d\mathbf{P}(\mathbf{r})/dt = \mathbf{j}(\mathbf{r})$ with eq. (3.3.10) and taking the matrix elements of its Fourier component, one obtains

$$\begin{aligned}i(E_{\lambda \mathbf{K}}/\hbar)(P_{-q})_{\lambda \mathbf{K},g} &= (-e/m) \left[\sum_i (1/2) \{ \exp(i\mathbf{q} \cdot \mathbf{r}_i) \mathbf{p}_i \right. \\ &\quad \left. + \mathbf{p}_i \exp(i\mathbf{q} \cdot \mathbf{r}_i) \} \right]_{\lambda \mathbf{K},g}.\end{aligned}\quad (8.4.2)$$

Note that $\Psi_{\lambda \mathbf{K}}$ is a linear combination of $\Psi_{\nu \mu \mathbf{k} \mathbf{K}}$ (see (8.3.2)) and that the matrix element of $[\cdots]$ between the Slater determinants $\Psi_{\nu \mu \mathbf{k} \mathbf{K}}$ and Φ_g representing a transition of one electron from $\mu \mathbf{k} - \mathbf{K}$ to $\nu \mathbf{k}$ can be written as (see also Appendix 1)

$$\int d\mathbf{r} \psi_{\nu \mathbf{k}}^*(\mathbf{r}) (1/2) \{ \exp(i\mathbf{q} \cdot \mathbf{r}) \mathbf{p} + \mathbf{p} \exp(i\mathbf{q} \cdot \mathbf{r}) \} \psi_{\mu \mathbf{k} - \mathbf{K}}(\mathbf{r}).$$

This is non-vanishing only when $\mathbf{K} = \mathbf{q}$ due to the Bloch theorem (7.1.2). Since q for visible light is much smaller than the reciprocal lattice, one can neglect the difference of the periodic function $u_{\mathbf{k}}(\mathbf{r})$ for small difference $\Delta\mathbf{k} = \mathbf{q}$. The above expression can then be approximated as

$$\delta_{\mathbf{K},\mathbf{q}} \mathbf{p}_{v\mu}(\mathbf{k}), \quad \text{where} \quad \mathbf{p}_{v\mu}(\mathbf{k}) \equiv \int d\mathbf{r} \psi_{v\mathbf{k}}^*(\mathbf{r}) (-i\hbar \nabla) \psi_{\mu\mathbf{k}}(\mathbf{r})$$

and the matrix element in (8.4.2) can be written as

$$[\cdot \cdot \cdot]_{\lambda\mathbf{K},\mathbf{g}} = \delta_{\mathbf{K},\mathbf{q}} \sum_{v\mu\mathbf{k}} \mathbf{p}_{v\mu}(\mathbf{k}) f_{\lambda\mathbf{K}}^*(v\mu\mathbf{k}). \quad (8.4.3)$$

If one neglects the electron-hole interaction in (8.3.3), and approximates $\mathbf{K} = \mathbf{q} \cong 0$, one can simply put $\lambda \rightarrow \mathbf{k}$, $E_{\lambda\mathbf{K}} = \varepsilon_v(\mathbf{k}) - \varepsilon_\mu(\mathbf{k})$, $f_{\lambda\mathbf{K}}(v\mu\mathbf{k}') = \delta_{\mathbf{k}'\mathbf{k}}$, and obtain from (8.4.1)

$$\text{Im}\chi_t(\omega) = \sum_{v\mu\mathbf{k}} (2e^2/m^2\omega^2) |p_{iv\mu}(\mathbf{k})|^2 \pi \delta[\{\varepsilon_v(\mathbf{k}) - \varepsilon_\mu(\mathbf{k})\} - \hbar\omega - i\gamma]. \quad (8.4.4)$$

This represents the absorption spectra due to the interband transition in the one-electron approximation. But for the \mathbf{k} dependence of $\mathbf{p}_{v\mu}(\mathbf{k})$, it reflects the joint density of states which rises as $[\hbar\omega - \varepsilon^{(m)}(0)]^{1/2}$ from the absorption edge $\varepsilon^{(m)}(0) = \varepsilon_g$ (see the broken line in Fig. 8.1(a)).

Let us now study the effect of the electron-hole Coulomb interaction, neglecting the ℓ -t splitting energy for the moment. If $F_{\lambda 0}^i(\mathbf{R})$ is a slowly varying function, the Fourier transform $f_{\lambda 0}(\mathbf{k})$ of $F_{\lambda 0}^i(\mathbf{R}) \exp(i\mathbf{k}_m(0) \cdot \mathbf{R})$ has a sharp peak at $\mathbf{k}_m(0)$ where one can use the expansion

$$\mathbf{p}_{v\mu}(\mathbf{k}) = \mathbf{p}_{v\mu}^{(m)} + (\mathbf{k} - \mathbf{k}_m(0)) \cdot \nabla \mathbf{p}_{v\mu}^{(m)} + \dots \quad (8.4.5)$$

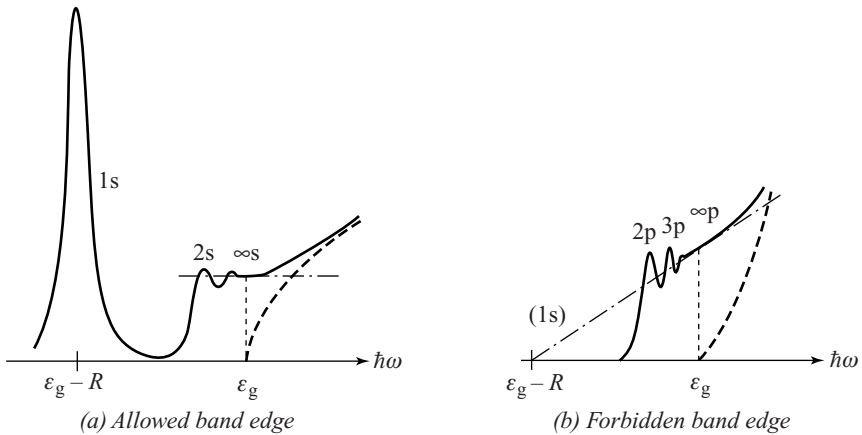


Fig. 8.1 The effect of the electron-hole Coulomb interaction on the fundamental absorption edge. The broken lines represent the interband absorption edge without the Coulomb effect.

to a good approximation. One can thus classify the following two cases with their characteristic spectral features, as first studied by Elliott.⁶

(a) *Allowed band edge*: $p_{v\mu}^{(m)} \neq 0$

One can neglect the second term of (8.4.5) and obtain

$$\text{Im}\chi_t(\omega) = \sum_{\lambda} (2N_0 e^2 / m^2 \omega^2) |p_{v\mu}^{(m)}|^2 |F_{\lambda}'(0)|^2 \pi \delta(E_{0\lambda 0} - \hbar\omega - i\gamma). \quad (8.4.6)$$

Only the s-like states with nonvanishing $F_{\lambda}'(0)$ appear in the absorption spectra as shown schematically by the solid line in Fig. 8.1(a). The *discrete* lines for bound states with $\lambda = n = 1, 2, \dots$ at

$$E_n = \varepsilon_g - Ry_{\text{ex}} n^{-2} \quad (8.4.7)$$

have relative intensities

$$|F_n'(0)|^2 = \pi^{-1} (v_0/a_{\text{ex}}^3) n^{-3}. \quad (8.4.8)$$

As n tends to ∞ , they overlap due to their finite spectral widths (to be studied later), forming a plateau of height

$$|F_n'(0)|^2 |dE_n/dn|^{-1} = (v_0/a_{\text{ex}}^3) (2\pi Ry_{\text{ex}})^{-1} \quad (8.4.9)$$

as shown in Fig. 8.1(a).

The s-like *continuous* states with radial component k of the wave vector \mathbf{k} for the relative motion with energy

$$E_k = \varepsilon_g + \hbar^2 k^2 / 2m_r \quad (8.4.10)$$

have relative amplitudes

$$|F_k'(0)|^2 = v_0(\pi\alpha) \exp(\pi\alpha) / \sinh(\pi\alpha) \quad (\alpha \equiv 1/ak) \quad (8.4.11)$$

due to the enhancement by Coulomb attraction. The density (per energy) of relative intensity is given by

$$\sum_k |F_k'(0)|^2 \delta(E_k - \hbar\omega). \quad (8.4.12)$$

With the use of state density $\delta(E_k - \hbar\omega) = (4\pi^2)^{-1} (2\mu/\hbar^2)^{3/2} E^{1/2}$, (8.4.12) is readily found to tend to the same value as (8.4.9) as $\hbar\omega$ tends to the continuum edge $E_0 = \varepsilon_g$, as shown in Fig. 8.1(a).

(b) *Forbidden band edge*: $p_{v\mu}^{(m)} = 0$

With the use of the second term of (8.4.5), one obtains the result with $p_{v\mu}^{(m)}$ replaced by $\nabla_{\mathbf{k}} p_{v\mu}^{(m)}$ and $F_{\lambda}'(0)$ by $\nabla_{\mathbf{R}} F_{\lambda}(0)$ in (8.4.6). As a result, only the p-like states appear in the spectra as shown schematically by the solid line in Fig. 8.1(b). Discrete lines

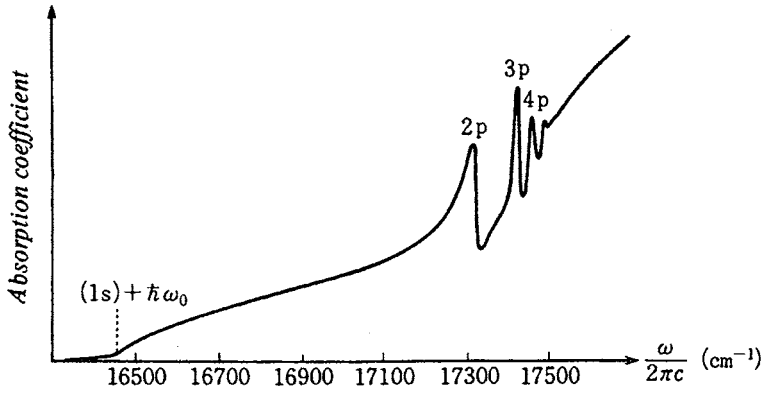


Fig. 8.3 The yellow series excitons at the fundamental absorption edge of Cu_2O crystal at 4.2 K. Due to Nikitine *et al.*⁸

dipole moment μ is provisionally given by

$$G_\lambda \equiv (2N_0\mu^2/\epsilon_{ce})|F'_\lambda(0)|^2. \quad (8.4.15)$$

Rewriting the valence-band index $\mu = p_x, p_y, p_z$ by $j = 1, 2, 3$ and the conduction-band index ν by s , making use of the relations

$$(e/m)(p_i)_{sj} = i\omega_{sj}e(x_i)_{sj}, \quad e(x_i)_{sj} = \mu\delta_{ij} \quad (8.4.16)$$

and approximating the transverse exciton energies $\hbar\omega_{\lambda t} = E_{0\lambda} - (1/3)G_\lambda$ by the bandgap energy $\epsilon_g = \hbar\omega_{sj}$, one can write (8.4.6) in a simpler form:

$$\text{Im}\chi_t(\omega)/\epsilon_{ce} = \sum_\lambda G_\lambda \pi \delta(\hbar\omega_{\lambda t} - \hbar\omega). \quad (8.4.17)$$

While the ℓ - t splitting G_λ with screening constant ϵ_{ce} (see (8.4.15)) contributed only from core electrons (see Fig. 5.2) plays the role of intensity in (8.4.17), a more consistent value G'_λ for the ℓ - t splitting will be obtained from the dispersion of $\epsilon(\omega)$.

Applying the dispersion relation (6.2.2) to $\text{Im}\alpha_t(\omega)$ which is an odd function, one readily obtains

$$\text{Re}\chi_t(\omega)/\epsilon_{ce} = \sum_\lambda 2\omega_{\lambda t}(G_\lambda/\hbar)(\omega_{\lambda t}^2 - \omega^2)^{-1}. \quad (8.4.18)$$

Remember that we are now studying the *valence* electrons contributing to dielectric constant $\epsilon(\omega)$ in a background medium with dielectric constant ϵ_{ce} . One can then apply eq. (6.3.2) relating dielectric constant $\epsilon(\omega)$ to susceptibility $\chi_t(\omega)$:

$$\epsilon(\omega) - \epsilon_{ce} = \chi_t(\omega). \quad (8.4.19)$$

Just as we saw in Chapters 5 and 6, poles $\omega_{\lambda t}$ and zeros $\omega_{\lambda \ell}$ of $\epsilon(\omega)$ which alternately appear as ω increases represent the energies of the transverse and longitudinal

excitons, respectively, which one can express in a compact formula:

$$\epsilon(\omega)/\epsilon_{\text{ce}} = \prod_{\lambda} (\omega_{\lambda\ell}^2 - \omega^2)/(\omega_{\lambda t}^2 - \omega^2), \quad (8.4.20)$$

similar to eq. (5.3.13). Under the assumption that $G_{\lambda} \ll \hbar\omega_{\lambda t}$, the energy $\hbar\omega_{\lambda\ell}$ of the longitudinal exciton is obtained by equating $\epsilon(\omega)$, given by (8.4.19) with (8.4.18), to zero. The thus-obtained ℓ - t splitting $G_{\lambda}^{\ell} \equiv \hbar(\omega_{\lambda\ell} - \omega_{\lambda t})$ is given by

$$G_{\lambda}^{\ell} \simeq (\epsilon_{\text{ce}}/\epsilon_{\lambda}^{\ell})G_{\lambda} = [2N_0\mu^2/\epsilon_{\lambda}^{\ell}]|F_{\lambda}^{\ell}(0)|^2, \quad (8.4.21)$$

$$\epsilon_{\lambda}^{\ell} \equiv \epsilon_{\lambda}^{\ell}(\omega_{\lambda\ell}), \quad (8.4.22)$$

$$\epsilon_{\lambda}^{\ell}(\omega) \equiv \epsilon_{\text{ce}} + \sum_{\lambda'(\neq\lambda)} 2\omega_{\lambda't} G_{\lambda'}^{\ell} / [\hbar(\omega_{\lambda't}^2 - \omega^2)]. \quad (8.4.23)$$

Namely, the ℓ - t splitting of the mode λ (and the corresponding dipole-dipole interaction) is screened by the *residual* dielectric constant $\epsilon_{\lambda}^{\ell}(\omega)$ defined by (8.4.23) at $\omega = \omega_{\lambda\ell}$ (which is to be sought for). Here “residual” means the dielectric constant with the contribution of the relevant mode λ being *deleted*. This is a reasonable result which can also be confirmed from the polariton dispersion (see Section 8.6).

It can be shown that the ratio of the ℓ - t splitting $G_{\lambda}^{\ell}(\lambda \gtrless n)$ to the energy separation from adjacent exciton lines, $\Delta E_n \equiv |dE_n/dn|$, is of the order of $(a_0/a_{\text{ex}})^2 \ll 1$ independent of n where a_0 is the lattice constant and a_{ex} is the radius of the 1s-exciton). This justifies a posteriori our procedure of treating ℓ - t splitting as a small quantity.

8.5 Spin-orbit and exchange interactions in a crystal field

The importance of the spin-orbit interaction rises rapidly as the atomic number of a constituent element of the material increases, as is well known in atomic physics. Let us consider the exciton with valence (and/or conduction) band originating from orbitally degenerate atomic states, and study how the spin-orbit interaction interplays with the exchange interaction.⁹ As a typical material, we consider this problem in alkali-halide crystals, in which the bottom of the conduction band is at $\mathbf{k} = 0$ with an orbitally non-degenerate Bloch function consisting mainly of the lowest unoccupied s state of the cation, and the top of the valence band is also at $\mathbf{k} = 0$ with a triply-degenerate (but for the spin-orbit interaction) Bloch function consisting of the outermost occupied p states of the anion.

The spin-orbit interaction $(\hbar/2m^2c_0^2)[\boldsymbol{\sigma} \times \nabla V(\mathbf{r})] \cdot \mathbf{p}$ can be approximated by $(\hbar^2/2m^2c_0^2)(dV/dr)\boldsymbol{\sigma} \cdot \mathbf{l}$ inside each anion since the periodic crystalline potential $V(\mathbf{r})$ becomes almost spherically symmetric in the neighborhood of the nucleus where the gradient of $V(\mathbf{r})$ takes the greatest value. Here $\hbar\mathbf{l} \equiv \mathbf{r} \times \mathbf{p}$ represents the orbital angular momentum of the electron around the anion nucleus. As in the case of an isolated atom, the triply-degenerate p-like valence bands at $\mathbf{k} = 0$ are split, due

to the spin-orbit interaction, into $j_v = 3/2$ and $j_v = 1/2$ states where $\mathbf{j}_v = \mathbf{l} + \boldsymbol{\sigma}$ is the total angular momentum of the electron. Since $2\boldsymbol{\sigma} \cdot \mathbf{l} = \mathbf{j}_v^2 - \mathbf{l}^2 - \boldsymbol{\sigma}^2 = j_v(j_v + 1) - 2 - 3/4$, the spin-orbit energies in the two states are given by $+\lambda/3$ and $-2\lambda/3$, where the splitting energy $\lambda(> 0)$ is expected to be close to the atomic value. However, the bottom of the conduction band is not split since $l = 0$ gives only $j_c = 1/2$.

In the j - j coupling scheme, the excited states $|j_c, j_v\rangle$ consists of $|1/2, 3/2\rangle$ and $|1/2, 1/2\rangle$, for which the square bracket in (8.3.11) with spin-orbit interaction included can be written as

$$\left[-(\hbar^2/2m_c)\nabla_\ell^2 - v(\mathbf{R}_\ell) + \varepsilon_g \right] + \begin{cases} -(1/3)\lambda (j_v = 3/2), \\ +(2/3)\lambda (j_v = 1/2). \end{cases} \quad (8.5.1)$$

Here ε_g denotes the bandgap energy in the absence of λ and the dispersion of the valence bands has been neglected. Since the sum of the total angular momentum of electron and hole is given by $\mathbf{J} = \mathbf{j}_c - \mathbf{j}_v$, one obtains $J = 2$ and 1 from $|1/2, 3/2\rangle$, and $J = 1$ and 0 from $|1/2, 1/2\rangle$. Among them, $J = 2$ and $J = 0$ states originate purely from spin triplet $S = 1$ since in the L - S coupling scheme ($\mathbf{J} = \mathbf{L} + \mathbf{S}$) one has $L = 1$ because $l_c = 0$ and $l_v = 1$. For these two states, one can use (8.5.1) as they stand, since the exchange interaction vanishes for the spin triplet.

On the other hand, the two sets of $J = 1$ states should be a linear combination of spin triplet and singlet. According to the well-known formula relating j - j and L - S coupling schemes, the $J = 1$ states obtained from $|1/2, 3/2\rangle$ and $|1/2, 1/2\rangle$ are given respectively by $(1/3)^{1/2}|\text{triplet}\rangle + (2/3)^{1/2}|\text{singlet}\rangle$ and $-(2/3)^{1/2}|\text{triplet}\rangle + (1/3)^{1/2}|\text{singlet}\rangle$. Calculating the exchange energy (8.3.18) with the use of 1s-state wave function $F(\mathbf{R}_\ell)$ of the purely orbital part $[\cdot \cdot \cdot]$ of (8.5.1) and adding the spin-orbit part, one obtains the first-order energy matrix

$$\begin{array}{l} \langle 1/2, 3/2 | \left[\begin{array}{cc} -(1/3)\lambda + (2/3)\Delta & (\sqrt{2}/3)\Delta \\ (\sqrt{2}/3)\Delta & +(2/3)\lambda + (1/3)\Delta \end{array} \right] \\ \langle 1/2, 1/2 | \end{array} \quad (8.5.2)$$

where the exchange energy Δ is given by

$$\Delta = 2|F(0)|^2 W(0). \quad (8.5.3)$$

To obtain $W(0)$, one puts, in (8.3.8), the s-like and p_x -like atomic functions ϕ_s and ϕ_x respectively for the Wannier functions a_v and a_μ , and takes the limit of $\mathbf{K} \rightarrow 0$ ($\mathbb{K} \perp x$) in (8.3.9) obtained therefrom, considering the optically-allowed transverse exciton. With the use of (8.4.22) for the effective screening

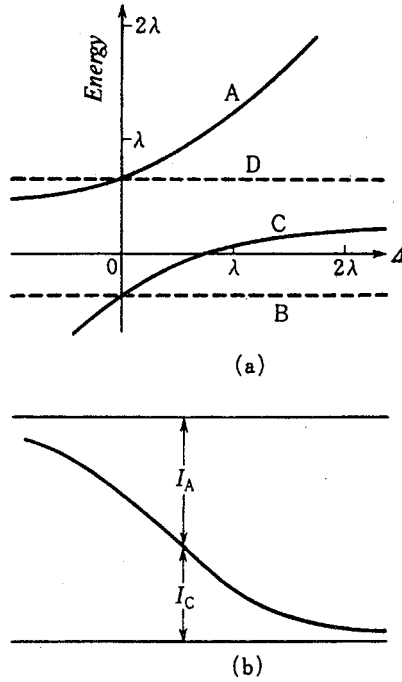


Fig. 8.4 (a) Schematic energy levels, with (b) their intensity ratio of absorption, of the 1s exciton in an alkali halide. Here λ represents the spin-orbit splitting of the valence band at the Γ point and Δ represents the electron-hole exchange energy.⁹

constant: $\epsilon' \equiv \epsilon'_{1s}(\omega_{1st})$, one obtains

$$W(0) = \iint d\mathbf{r} d\mathbf{r}' \phi_s(\mathbf{r}) \phi_x(\mathbf{r}) [e^2 / 4\pi \epsilon_0 |\mathbf{r} - \mathbf{r}'|] \phi_x(\mathbf{r}') \phi_s(\mathbf{r}') - N_0 \mu^2 / 3\epsilon'. \quad (8.5.4)$$

The eigenvalues of the energy matrix (8.5.2) for $J = 1$, as functions of Δ , are shown by solid lines C and A in Fig. 8.4(a), together with the Δ -independent energies of the pure triplets $J = 2$ and 0 shown by broken lines B and D. In Fig. 8.4 (b) is shown the partition of the oscillator strengths between C and A which is obtained from the weights of the optically-allowed spin singlet state contained in the eigenstates of (8.5.2). With $\Delta = 0$ ($j-j$ coupling), the intensity ratio is given by $C(j_v = 3/2) : A(j_v = 1/2) = 2 : 1$ as a reflection of multiplicity $2j_v + 1$, while it tends to $0 : 1$ in the $L-S$ limit of $\Delta \rightarrow \infty$ (or $\lambda \rightarrow 0$) where C and A tend to pure triplet and singlet states, respectively.

The deviation of the intensity ratio of the spin-orbit components $\Gamma(1/2, 3/2)$ and $\Gamma(1/2, 1/2)$ from 2:1, as shown in Fig. 8.4 and in fact found in the observed spectra of Fig. 8.2, is due to the exchange interaction; conversely, one can estimate Δ and λ from this deviation and the splitting energy.

A continuous variation in the intensity ratio of the spin-orbit split lines as a function of λ/Δ has been observed by Kato, Yu and Goto¹⁰ in a mixed crystal system: $\text{CuCl}_{1-x}\text{Br}_x$ ($0 < x < 1$). Both CuCl and CuBr of the zinc-blende type have band structures similar to those of alkali halides in that the bottom of the conduction band and the top of the valence band are at $k = 0$, respectively, with single and triple (but for the spin-orbit coupling) orbital degeneracies. However, the valence band consists not only of halide p-orbitals but of also Cu d-orbitals. Depending sensitively on this mixing ratio, the energy difference λ of the spin-orbit bands $Z_{1,2}(j_v = 3/2)$ and $Z_3(j_v = 1/2)$ is negative in CuCl while positive in CuBr. Since the exchange energy Δ given by (8.5.3) is small compared to $|\lambda|$ in these crystals due to their excitonic radii being larger than those in alkali halides, the situation is closer to the $j-j$ coupling scheme so that the intensity ratio of the first and the second exciton peaks is not far from 1:2 in CuCl and 2:1 in CuBr.

It is known that the mixed crystals $\text{CuCl}_{1-x}\text{Br}_x$ are also of zinc blende structure in the entire concentration range: $0 < x < 1$. The observed peak positions and intensity ratio of the first and second exciton peaks are shown in Fig. 8.5 (a) and (b),

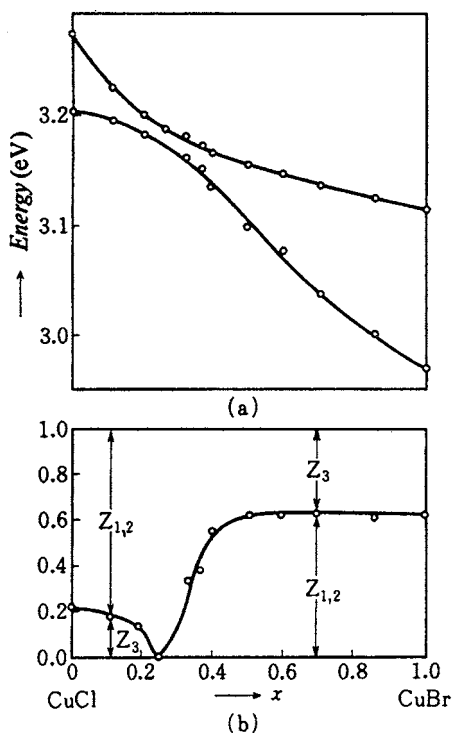


Fig. 8.5 (a) Observed energies, and (b) intensity ratio of $Z_{1,2}$ and Z_3 excitons in a mixed crystal: $\text{CuCl}_{1-x}\text{Br}_x$ as functions of the concentration x . Due to Kat., Yu and Goto.¹⁰

respectively, as functions of the concentration x . Assuming that the mixed crystals have band structures determined by the averaged periodic potential of the two components, one would expect that λ varies continuously (a band structure of the amalgamation type to be discussed in Section 8.8) from negative to positive values as x increases from 0 to 1, but that Δ , being always positive, does not vary significantly. The ratio Δ/λ would vary in such a way that in Fig. 8.4(a) one proceeds to the left starting from a slightly negative value of the abscissa, turns from $-\infty(\lambda = -0)$ to $+\infty(\lambda = +0)$, proceeds again to the left down to a slightly positive value. For $\lambda < 0$, one must invert the ordinate of Fig. 8.4, and hence transpose the first and the second peaks. In this way, one sees how the characteristic features of the observation shown in Fig. 8.5 can be explained in terms of the competition of the spin-orbit and exchange interactions.

So far we have confined ourselves to situations which can be described within the atomic model for the interplay of exchange and spin-orbit interactions. We will now describe the general situations with the effects of crystalline field and external field taken into account. The valence band in ZnO, due to the uniaxial crystal field of this wurtzite type structure, is split into three sheets A, B and C (not to be confused with A, B, C and D of Fig. 8.4 which are two-particle states) which should not split any more under perturbation with time-reversal symmetry since each of them is a Kramers doublet. In spite of this, Koda and Langer¹¹ found that each of the A and B exciton lines is split into a doublet under uniaxial stress along the c -axis, and ascribed this to the exciton effect. In fact, Akimoto and Hasegawa¹² explained this in terms of the electron-hole exchange interaction under stress, and estimated therefrom the exchange energy.

Koda *et al.*¹³ also found that a uniaxial stress applied on a CuCl crystal gives rise to significant mixing of the optically-allowed transverse exciton and the forbidden longitudinal exciton, and ascribed this to the k -linear term in the valence-band structure caused by the stress in this crystal without inversion symmetry.

A systematic formulation to describe more general situations of orbital and spin degeneracies of the conduction and valence bands taking into account the crystal fields and external fields has been presented by Cho¹⁴ and applied to the analysis of exciton spectra in various semiconductors (see also Section 11.2).

8.6 Polaritons and optical constants

The optical transition for creating an exciton or an optical phonon, the typical elementary excitation associated with electric polarization, has *enormous* oscillator strength of the order of the number N_0 of atoms included in a solid of macroscopic size under consideration, which is concentrated in a single absorption line. This

is obvious from (8.4.6)), or more generally from the sum rule (6.2.5), the r.h.s. summation of which extends over N_0 particles. It may be inappropriate to describe such a strong absorption process by a simple perturbation theory. In fact, in matter, the electromagnetic waves with a frequency ω in the vicinity of the intrinsic absorption region are so strongly coupled with the (transverse) polarization waves that it is inappropriate to treat them as separate entities with only a one-way transfer of energy. On the contrary, one must consider a coupled wave of the electromagnetic and polarization fields (energy is transferred so to speak *to and fro* between them; namely, absorption and emission take place alternately)¹⁵ before describing the optical absorption process. The quantum of the coupled wave is called a *polariton* with the implication of being partly a photon and partly a quantum of polarization (exciton or optical phonon), after Hopfield¹⁶ who first presented a quantum-mechanical description of the interacting Boson fields of photon and exciton. In fact, the corpuscular nature of the polariton reveals itself in a variety of optical processes, as will be seen later. Here, however, we will start with the classical field description of the coupled waves since this is more transparent in relating them to the dielectric dispersion which is a vital aspect of the polariton picture.

Let us confine ourselves to isotropic solids, with dielectric constant $\epsilon(\omega)$ and magnetic permeability μ_0 which is the same as in a vacuum. The equation for the transverse part of the electromagnetic field within the matter, obtained from eqs. (1.1.1, 1.1.2, 1.1.6) with $\mathbf{J} = 0$, reads:

$$\mu_0^{-1} \nabla^2 \mathbf{E} = \partial^2 \mathbf{D} / \partial t^2. \quad (8.6.1)$$

Putting in a plane-wave solution: $\exp(i\mathbf{k} \cdot \mathbf{r} - i\omega t)$, dividing by $-\omega^2 \epsilon_0 \mathbf{E}$ and making use of (1.1.5), one obtains the dispersion

$$c_0^2 k^2 / \omega^2 = \epsilon(\omega) / \epsilon_0 = n(\omega)^2, \quad (8.6.2)$$

where $c_0 = (\epsilon_0 \mu_0)^{-1/2}$ is the velocity of light in a vacuum and $n(\omega)$ is the refractive index of the matter. The phase velocity of light in matter is given by

$$c(\omega) = \omega / k = c_0 / n(\omega). \quad (8.6.3)$$

The electric polarization of matter consists of contributions from various modes of oscillators, which may be atomic (optical modes of lattice vibrations) or electronic (excitons). Out of them we consider explicitly a particular oscillator with its contribution \mathbf{P} , confining ourselves to the region of ω in the vicinity of ω_t , the eigenfrequency of the oscillator. Describing the contributions from all higher-frequency oscillators phenomenologically by the *residual* dielectric constant (see Section 8.4), ϵ' (assumed here to be positive), while neglecting (legitimately) those

from lower-frequency ($\omega_t \ll \omega$) oscillators (if any), one can write

$$D = \epsilon' E + P. \quad (8.6.4)$$

At frequency ω significantly smaller than ω_t , P will follow E instantaneously with polarizability χ of *that* oscillator. For example, χ is given by $N_0 e^2 / M_r \omega_t^2$ in the case of optical lattice vibration in a diatomic crystal (see (5.3.14)), and by $4N_0 \mu^2 |F_{1s}(0)|^2 / E_{01s}$ with $\omega_t \cong E_{01s} / \hbar$ (see (8.4.15, 8.4.18)) in the case of a 1s Wannier exciton. The equation of motion for this oscillator is then given by

$$\partial^2 P / \partial t^2 + \omega_t^2 P = \omega_t^2 \chi E. \quad (8.6.5)$$

On the other hand, the electromagnetic equation (8.6.1), combined with (8.6.4), can be written as

$$\epsilon' \partial^2 E / \partial t^2 + \epsilon_0 c_0^2 k^2 E = -\partial^2 P / \partial t^2. \quad (8.6.6)$$

Since $P \parallel E$ due to isotropy, the plane-wave solutions of the coupled equations (8.4.5, 8.4.6) satisfy

$$\begin{pmatrix} -\omega^2 + \omega_t^2 & -\chi \omega_t^2 \\ -\omega^2 & -\epsilon' \omega^2 + \epsilon_0 c_0^2 k^2 \end{pmatrix} \begin{pmatrix} P \\ E \end{pmatrix} = 0. \quad (8.6.7)$$

By equating the determinant of the coefficients to zero, one obtains the eigenfrequencies ω of the P - E coupled modes as functions of wave vector k . Alternatively, one can solve for k as a function of ω :

$$\begin{aligned} c_0^2 k^2 / \omega^2 &= \epsilon(\omega) / \epsilon_0 = [\epsilon' + \chi \omega_t^2 / (\omega_t^2 - \omega^2)] / \epsilon_0 \\ &= (\epsilon' / \epsilon_0) (\omega_t^2 - \omega^2) / (\omega_t^2 - \omega^2), \end{aligned} \quad (8.6.8)$$

$$\omega_t = \omega + \Delta, \quad \Delta = [(1 + \chi / \epsilon')^{1/2} - 1] \omega_t. \quad (8.6.9)$$

The energy ϵ' may be assumed to be ω independent if the eigenfrequencies of the other oscillators are far enough away from the region $\omega \sim \omega_t$ with which we shall be concerned.

Equation (8.6.8) gives the wave vector k and dielectric constant ϵ as functions of ω , which are schematically shown in Fig. 8.6(a), (b). Namely, the electromagnetic wave propagating through matter with dielectric dispersion as given in Fig. 8.6(b) has the $\omega(k)$ dispersion as shown in Fig. 8.6(a). The quantum of this coupled wave is called the polariton, as mentioned already. It is to be noted that there are two branches of polariton for a given k .

It does not make sense to ask whether the electromagnetic field of the polariton comes from the external source or not. The polarization field P is always accompanied by the electric field E and vice versa, through the coupled equations (8.6.5, 8.6.6), irrespective of the origin of E . Denoting the intersecting point of the

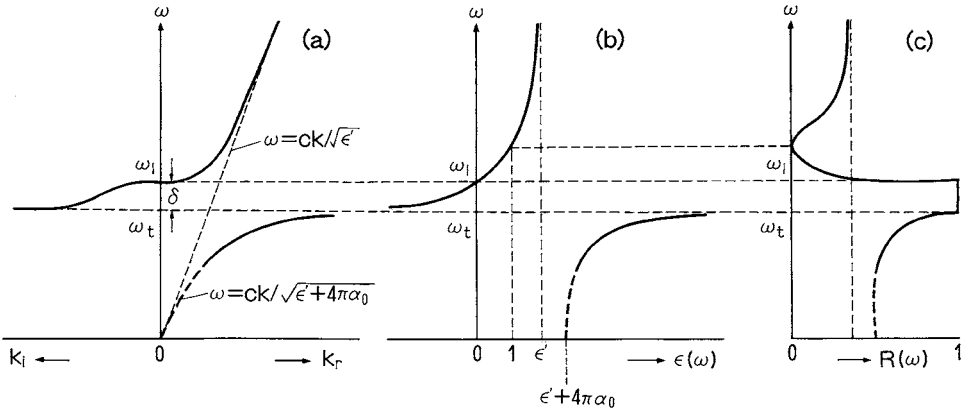


Fig. 8.6 (a) Dispersion of polariton, $\omega(k)$. (b) Dielectric dispersion, $\epsilon(\omega)$. (c) Reflectivity, $R(\omega)$.

uncoupled polarization wave and electromagnetic wave (shown by dotted lines in Fig. 8.6(a)) by $k_p \equiv (\epsilon')^{1/2}\omega/c_0$, one finds that the lower and upper branches of the polariton in region (I) $k \gg k_p$ are respectively a pure polarization wave and a renormalized electromagnetic wave with index of refraction $n = (\epsilon')^{1/2}$. In region (III) $k \ll k_p$, the opposite limit, the lower branch represents a renormalized electromagnetic wave with index of refraction $n = (\epsilon' + \chi)^{1/2}$, the difference from the upper branch of region (I) being the participation of the particular oscillator of our concern. The upper-branch polariton in this region is subject to the depolarizing field $\mathbf{E} = -\mathbf{P}/\epsilon'$ since the second term on the l.h.s. of (8.6.6) is negligible. This is the reason why the frequency of the upper-branch polariton tends to the longitudinal frequency ω_l with vanishing k , in spite of its transverse nature. The intermediate region (II) around k_p is the typical polariton region where the continuous change of the two features – electromagnetic and matter waves – takes place within each branch. Simple notions of elementary processes, such as absorption and emission of a photon by matter, must be reconsidered more carefully as will be done below and in later sections.

For the reasons mentioned above, the spectroscopic observation of the polarization wave in matter reduces to the problem of a boundary condition through which the polariton inside the matter is connected to the electromagnetic wave outside. An electromagnetic wave of the form: $\exp(ik_0z - i\omega t)$ incident perpendicularly upon the surface of the matter will give rise to a polariton wave of the form: $\exp[ik(\omega)z - i\omega t]$ inside the matter, where k_0 and $k(\omega)$ are given by ω/c_0 and $[\epsilon(\omega)]^{1/2}\omega/c_0$, respectively (see (8.6.8)). The latter is shown schematically in Fig. 8.6(a). The fraction R of incident electromagnetic energy which is reflected by the surface and the fraction $1 - R$ which penetrates into the matter as a polariton

are determined by the boundary conditions associated with Maxwell's equation – the continuity, across the surface, of the tangential components of \mathbf{E} and \mathbf{H} and normal components of \mathbf{D} and \mathbf{B} . For normal incidence, the reflectivity is given by

$$R(\omega) = |[n(\omega) - 1]/[n(\omega) + 1]|^2 = [(n_r - 1)^2 + n_i^2]/[(n_r + 1)^2 + n_i^2], \quad (8.6.10)$$

where $n(\omega) = n_r(\omega) + in_i(\omega) = \epsilon(\omega)^{1/2}$ is the complex refractive index. The reflectivity, $R(\omega)$, corresponding to Fig. 8.6(a), (b) is shown in Fig. 8.6(c).

In the region $\omega_t < \omega < \omega_\ell$, the dielectric constant is negative while the refractive index and the wave vector is purely imaginary as shown in the left halves of Fig. 8.6 (a), (b). There is no propagating mode in this ω region. Instead, \mathbf{E} and \mathbf{P} have the form $\exp[-k_i(\omega)z - i\omega t]$, with amplitude decaying exponentially with *attenuation coefficient* $k_i(\omega)$. This does not mean, however, that the incident electromagnetic energy is absorbed by the matter. Instead, it is totally reflected by the surface without energy loss as shown in Fig. 8.6 (c). The imaginary part of $\epsilon(\omega)$ which is related to the absorption is zero in the present case, although $n(\omega)$ and $k(\omega)$ are purely imaginary.

However, the electromagnetic wave with $\omega = \omega_t$ should be subject to the resonant absorption: $\epsilon(\omega)$ with real part (8.6.8) should have the imaginary part of $\delta(\omega - \omega_t)$ type according to the dispersion relation (6.2.2, 6.2.3). To be more realistic, one can take into account a damping of the oscillator phenomenologically by adding the term $-\gamma dP/dt$ to the r.h.s. of (8.6.5). This modifies the dielectric constant (8.6.8) into

$$\epsilon(\omega) = \epsilon' + \chi\omega_t^2/[\omega_t^2 - \omega^2 - i\gamma\omega]. \quad (8.6.11)$$

The situation varies depending on the relative magnitude of the damping constant γ and the ℓ -t splitting Δ given by (8.6.9). In the “dissipative” case where

$$\Delta \ll \gamma/2 \ll \omega_t, \quad (8.6.12)$$

one finds from (8.6.9) that $\chi/2\epsilon' \sim \Delta/\omega_t < 1$ and, hence, that the second term of (8.6.11) is small compared to the first by a factor $\Delta/(\gamma/2)$ even at resonance: $\omega = \omega_t$. One then finds from (8.6.2) that the *energy density* of the electromagnetic wave decays exponentially as it penetrates the matter, with the absorption coefficient (twice the attenuation coefficient for the *amplitude*):

$$\begin{aligned} A(\omega) &= 2k_i(\omega) \cong (\omega/c_0)(\epsilon')^{-1/2} \text{Im}\{\epsilon(\omega)\} \\ &\cong (\omega/c_0)(\epsilon')^{1/2} \Delta(\gamma/2)/[(\omega_t - \omega)^2 + (\gamma/2)^2]. \end{aligned} \quad (8.6.13)$$

Namely, the spatial attenuation of an electromagnetic wave in this case is almost entirely due to the absorption of its energy by the matter. Within the narrow resonance region in which ϵ' can be taken as constant, it is proportional to $\omega \text{Im}\epsilon(\omega)$. The lineshape of the absorption spectrum, except for the factor ω , is a Lorentzian, with full width at half-maximum being given by the damping constant γ .

On the other hand, it is in the “dispersive” case:

$$\gamma/2 \ll \Delta \quad (8.6.14)$$

that the polariton effect manifests itself most clearly. In the region $\gamma/2 < \omega - \omega_t < \Delta$, namely, in the greatest part of the ℓ -t gap ($\omega_t < \omega < \omega_\ell$), the spatial attenuation of the electromagnetic wave is associated with total reflection (see Fig. 8.6(c)), not with energy absorption by the matter.

In the spectroscopic study of matter, it is more usual to measure (twice) the attenuation coefficient which is directly related with the energy absorption. This is justified only for cases of small oscillator strength such as impurity absorption bands and weak host absorption lines. For strong exciton lines such as in alkali halides, the conductivity spectra, $\text{Re}\{\sigma(\omega)\} = \omega \text{Im}\{\epsilon(\omega)\}$, obtained by the Kramers–Kronig analysis of the measured reflectivity spectra, has been preferred (to the absorption spectra) as the more fundamental quantity related with energy absorption as was shown in Fig. 8.2.⁷

The notion of coupled modes of electromagnetic and polarization waves was known with optical phonons,¹⁵ before the word polariton was introduced for the case of excitons. At present, one distinguishes them by “phonon polariton” and “exciton polariton”. We will have many occasions of referring to the latter in connection with a variety of excitonic processes.

8.7 Coherent potential approximation for exciton in mixed crystals

Some crystals of chemically similar species can often form binary mixed crystals or alloys with the same crystal structures as the pure ones over the entire range of their mixing ratio x ($0 < x < 1$). As will be shown in Section 8.9, the observed exciton spectra of such binary mixed crystals can be classified into two typical cases. In the *amalgamation* (A) type, there is *one spectral peak* of the exciton state with the same quantum number, whose *energy position* varies continuously with x , as seen in alkali halides with mixed alkalis such as $\text{K}_x\text{Rb}_{1-x}\text{Cl}$. In the *persistence* (P) type, there appear *two spectral peaks* of the same exciton state corresponding to the two constituent crystals, whose *intensity ratio* varies continuously with x , as in alkali halides with mixed halides such as $\text{KCl}_{1-x}\text{Br}_x$. On the other hand, $\text{CuCl}_{1-x}\text{Br}_x$ shown in Fig. 8.5 belongs to the amalgamation type in spite of being halide mixtures: each of the spin–orbit split components, $Z_{1,2}$ and Z_3 peaks, moves continuously and almost linearly with x except in the anomalous region of anticrossing due to the exchange interaction.

As the simplest model, one can characterize an exciton in a mixed crystal by two parameters: Δ and B . Here Δ is the energy difference of the exciton peaks in the two constituent crystals (to be distinguished from the exchange energy Δ of Section 8.5)

while $2B$ is the total width of the exciton band assumed to be common to the two. It is obvious that experimentally classified A and P types correspond respectively to the limits of $\Delta/B \ll 1$ and $\gg 1$ of this model. The exciton moving from site to site with transfer energy $|t| \sim B/\nu$ (where ν is the number of nearest neighbors in the crystal) is subject to a potential fluctuation of amplitude Δ depending on the atomic species of the site. If $\Delta \ll B$, the exciton will move in the averaged potential, being weakly scattered by the fluctuation which results in a small level width and shift. If $\Delta \gg B$, the exciton will prefer to be localized in either site than moving about. Therefore, the persistence versus amalgamation problem is a special case of the *localization versus delocalization* problem of a particle moving in spatially fluctuating fields.

In order to see how the spectral features change from the A type to the P type as Δ/B increases, one must find a theory which works somehow throughout the intermediate region. For this purpose, a heuristic method interpolating the self-energy of the moving particle between the two limits was proposed and applied to the calculation of exciton spectra,¹⁷ as shown below. The method itself turned out to be equivalent to the *coherent potential approximation* (CPA) developed to study lattice vibrations¹⁸ and the electronic structures of alloys,¹⁹ as was most clearly shown by Velicky *et al.*²⁰ We will present here the CPA formalism in view of its generality.

Let us begin with the Frenkel exciton in a pure crystal described in Section 8.1, and denote the state Φ_n of (8.1.1), in which only the n th site atom is excited, by the ket vector $|n\rangle$. The Hamiltonian for a Frenkel exciton can then be written as

$$\mathbb{H} = \sum_n |n\rangle E_n \langle n| + \sum_n \sum_m |n\rangle t_{nm} \langle m| \quad (t_{nn} = 0). \quad (8.7.1)$$

Here the diagonal element E_n represents the atomic excitation energy at the n th site which, in Section 8.1, was denoted by ε independent of n due to the translational symmetry of the crystal lattice. For the same reason, the non-diagonal element t_{nm} , the transfer energy of the exciton between the n th and m th sites, depends only on the difference $\mathbf{R}_n - \mathbf{R}_m$ of their positions as is explicitly given by (8.1.6) for the Frenkel exciton, and the eigenstates $|\mathbf{K}\rangle$ of (8.7.1) is given by the running wave type linear combination of $|n\rangle$ as given by (8.1.7). The print-type \mathbb{H} is used for the operators (matrices) to distinguish them from ordinary numbers, only for the sake of mathematical clarity in the following description.

For a binary mixed crystal consisting of atoms (or molecules) A and B with respective concentrations $c_A \equiv c$ and $c_B = 1 - c$ randomly distributed over the regular lattice sites, we assume that the intrasite excitation energy E_n takes a value E_A or E_B according as the site is occupied by atom A or B, but the transfer energies t_{nm} take common values irrespective of the species of atoms occupying the sites n and m , just for simplicity (the so-called *diagonal randomness model*). Let us

now decompose \mathcal{H} to its ensemble average $\mathcal{H}_0 \equiv \langle\langle \mathcal{H} \rangle\rangle$ and the fluctuation \mathcal{V} therefrom:

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{V}, \quad (8.7.2)$$

$$\mathcal{H}_0 = \sum_n |n\rangle \bar{E} \langle n| + \sum_n \sum_m |n\rangle t_{nm} \langle m| = \sum_{\mathbf{K}} |\mathbf{K}\rangle E_{\mathbf{K}} \langle \mathbf{K}|, \quad (8.7.3)$$

$$\mathcal{V} = \sum_n |n\rangle v_n \langle n| \equiv \sum_n \mathcal{V}_n, \quad v_n = E_n - \bar{E}, \quad (8.7.4)$$

$$\bar{E} \equiv \langle\langle E_n \rangle\rangle = c_A E_A + c_B E_B, \quad (8.7.5)$$

$$E_{\mathbf{K}} = \bar{E} + t_{\mathbf{K}}, \quad t_{\mathbf{K}} = \sum_n t_{0n} \exp(i\mathbf{K} \cdot \mathbf{R}_n). \quad (8.7.6)$$

Because of the random occupation of the sites by A and B atoms, the intersite correlation can be assumed to vanish:

$$\langle\langle v_n v_m \rangle\rangle = \langle\langle v_n \rangle\rangle \langle\langle v_m \rangle\rangle = 0. \quad (8.7.7)$$

The principle of the CPA is to incorporate into \mathcal{H}_0 the *main* effect of exciton scattering by \mathcal{V} in the form of an *average effective potential* u (or *self-energy* in a different terminology), in such a way that the residual effect of scattering by the remaining potential

$$\mathcal{V}' \equiv \mathcal{V} - u\mathbb{I} = \sum_n |n\rangle v'_n \langle n| \equiv \sum_n \mathcal{V}'_n, \quad v'_n \equiv v_n - u, \quad (8.7.8)$$

$$\text{where } \mathbb{I} \equiv \sum_n |n\rangle \langle n| = \sum_{\mathbf{K}} |\mathbf{K}\rangle \langle \mathbf{K}|, \quad (8.7.9)$$

is *minimized* (in a way to be seen below). Here \mathbb{I} is a *unit* (or *identity*) operator *within the subspace* of one Frenkel exciton. Let us define the *resolvent* operator (sign reversed from the conventional definition) by

$$\mathcal{R}(z) \equiv (z\mathbb{I} - \mathcal{H})^{-1}, \quad (8.7.10)$$

where $z \equiv E - i\gamma$ is to be understood as a complex number with infinitesimally small negative imaginary part ($\gamma \rightarrow +0$). If all the eigenvalues E_j and eigenvectors $|j\rangle$ of \mathcal{H} are obtained, the state density of the exciton is given by

$$\rho(E) \equiv \sum_j \delta(E - E_j) = \pi^{-1} \text{Im} \left[\sum_j \langle j | \mathcal{R}(z) | j \rangle \right] \equiv \pi^{-1} \text{Im} [\text{Tr} \mathcal{R}(z)]. \quad (8.7.11)$$

The optical absorption spectrum is given by

$$F(E) = \sum_j |\langle j | \mathbf{K} = \mathbf{0} \rangle|^2 \delta(E - E_j) = \pi^{-1} \text{Im} [\langle \mathbf{0} | \mathcal{R}(z) | \mathbf{0} \rangle], \quad (8.7.12)$$

since only the $\mathbf{K} = \mathbf{0}$ component of the state $|j\rangle$ contributes to the optical transition (see Section 8.4) provided that the transition dipole moments of A and B atoms are equal. The quantities of physical importance are thus derivable from the resolvent operator $\mathcal{R}(z)$.

In order to calculate $\mathbb{R}(z)$, we expand it in powers of V' (instead of V):

$$\begin{aligned}\mathbb{R}(z) &= [z\mathbb{I} - (H_0 + u\mathbb{I}) - V']^{-1} \\ &= \mathcal{D} + \mathcal{D} \sum_n V'_n \mathcal{D} + \mathcal{D} \sum_n V'_n \mathcal{D} \sum_m V'_m \mathcal{D} + \cdots, \quad (8.7.13)\end{aligned}$$

where \mathcal{D} is an operator diagonal in \mathbf{K} :

$$\mathcal{D}(z) \equiv [z\mathbb{I} - (H_0 + u\mathbb{I})]^{-1} = \sum_{\mathbf{K}} |\mathbf{K}\rangle (z - E_{\mathbf{K}} - u)^{-1} \langle \mathbf{K}|. \quad (8.7.14)$$

A part of the expansion (8.7.13) can be rearranged in such a way that successive scattering at the same site are collected into the so-called t -matrix, which we write here by upper case T to distinguish it from the transfer energy t appearing in (8.7.1)–(8.7.6):

$$\begin{aligned}V'_n + V'_n \mathcal{D} V'_n + V'_n \mathcal{D} V'_n \mathcal{D} V'_n + \cdots &= |n\rangle v'_n (1 - G v'_n)^{-1} \langle n| \\ &\equiv |n\rangle T_n \langle n| \equiv T_n, \quad (8.7.15)\end{aligned}$$

$$G(z) \equiv \mathcal{D}(z)_{nn} = N^{-1} \sum_{\mathbf{K}} (z - E_{\mathbf{K}} - u)^{-1} = G^{(0)}(z - u), \quad (8.7.16)$$

$$G^{(0)}(z) \equiv \mathcal{D}^{(0)}(z)_{nn} = N^{-1} \sum_{\mathbf{K}} (z - E_{\mathbf{K}})^{-1}. \quad (8.7.17)$$

Note that the imaginary part of $G^{(0)}(z)$ represents π times the normalized density of states $\rho^{(0)}(E)$ in the unperturbed lattice. The expansion (8.7.13) can then be rewritten as

$$\begin{aligned}\mathbb{R}(z) &= \mathcal{D} + \mathcal{D} \sum_n T_n \mathcal{D} + \mathcal{D} \sum_n T_n \mathcal{D} \sum_{m(\neq n)} T_m \mathcal{D} \\ &\quad + \mathcal{D} \sum_n T_n \mathcal{D} \sum_{m(\neq n)} T_m \mathcal{D} \sum_{\ell(\neq m)} T_\ell \mathcal{D} + \cdots \quad (8.7.18)\end{aligned}$$

where none of the successive pairs of T scattering takes place at the same site. Then, in view of (8.7.7), it would be a good approximation to replace the ensemble average of the product of T s by the product of the averages of T :

$$\begin{aligned}\langle \langle \mathbb{R}(z) \rangle \rangle &\simeq \mathcal{D} + \mathcal{D} \sum_n \langle \langle T_n \rangle \rangle \mathcal{D} + \mathcal{D} \sum_n \langle \langle T_n \rangle \rangle \mathcal{D} \sum_{m(\neq n)} \langle \langle T_m \rangle \rangle \mathcal{D} \\ &\quad + \mathcal{D} \sum_n \langle \langle T_n \rangle \rangle \mathcal{D} \sum_{m(\neq n)} \langle \langle T_m \rangle \rangle \mathcal{D} \sum_{\ell(\neq m)} \langle \langle T_\ell \rangle \rangle \mathcal{D} + \cdots. \quad (8.7.19)\end{aligned}$$

In fact, the lowest-order discrepancy appears in the third-order terms:

$$\langle \langle T_n T_m T_\ell \rangle \rangle \neq \langle \langle T_n \rangle \rangle \langle \langle T_m \rangle \rangle \langle \langle T_\ell \rangle \rangle \text{ when } \ell = n.$$

Let us choose the effective potential u in such a way that the ensemble average of the T -matrix vanishes:

$$\langle \langle T_n \rangle \rangle = \langle \langle v'_n (1 - G v'_n)^{-1} \rangle \rangle = 0, \quad (8.7.20)$$

$$\text{namely, } G(z) = \langle \langle [G(z)^{-1} - v_n + u]^{-1} \rangle \rangle. \quad (8.7.21)$$

The coherent potential $u(z)$ as well as the individual potential $v_n^i(z)$ thus determined are energy-dependent complex quantities. Equation (8.7.19) is then written, with the use of (8.7.14), as

$$\langle\langle \mathcal{R}(z) \rangle\rangle \cong \mathcal{D}(z) = [z\mathbb{I} - \{\mathcal{H}_0 + u(z)\mathbb{I}\}]^{-1} = \sum_{\mathbf{K}} |\mathbf{K}\rangle [z - E_{\mathbf{K}} - u(z)]^{-1} \langle \mathbf{K} |. \quad (8.7.22)$$

The approximation (8.7.22) with $u(z)$ determined by (8.7.20) or (8.7.21) is the CPA. The normalized state density (8.7.11) and the optical spectrum (8.7.12) are then given by

$$\rho(E) \cong \pi^{-1} \text{Im}[G^{(0)}(E - u(z))], \quad (8.7.23)$$

$$F(E) \cong \pi^{-1} \text{Im}[\{z - E_0 - u(z)\}^{-1}], \quad (8.7.24)$$

where use has been made of (8.7.16, 8.7.17).

The CPA is exact in various limiting situations, as that was the primary aim of the heuristic method. In the weak scattering regime, it agrees with the perturbation theory up to the second order as is obvious from the statement which follows (8.7.19). Remembering (8.7.4 and 8.7.8), one obtains from (8.7.21)

$$u(z) = \langle\langle v_n^2 \rangle\rangle G(z) + (\text{higher-order terms in } v), \quad (8.7.25)$$

$$\langle\langle v_n^2 \rangle\rangle = c(1 - c)\Delta^2, \quad \Delta \equiv E_B - E_A. \quad (8.7.26)$$

Namely, (8.7.25) reduces, in its lowest order, to the *renormalized* ($G(z)$ instead of $G^{(0)}(z)$) second-order perturbation theory.

The CPA is exact again in the strong scattering limit: $\Delta \gg B$. Neglecting the band width, one can put $E_K \cong \bar{E}$ and $G^{(0)}(z) \cong (z - \bar{E})^{-1}$. Because of (8.7.16), (8.7.21) can be written as

$$[z - \bar{E} - u(z)]^{-1} \cong \langle\langle [z - \bar{E} - v_n]^{-1} \rangle\rangle. \quad (8.7.27)$$

Hence, the normalized density of states and the optical absorption spectrum reduce to

$$\rho(E) \simeq F(E) \cong c_A \delta(E - E_A) + c_B \delta(E - E_B). \quad (8.7.28)$$

8.8 Giant oscillator strength of an impurity-bound exciton

The coherent potential approximation stated in Section 8.7 turns out to be exact also in the dilution limit ($c \equiv c_A \rightarrow 0$ or $c_B \rightarrow 0$), as is obvious from the fact that single site scattering is considered up to infinite order as given by (8.7.15). Let us rewrite (8.7.21) as

$$u(z) - v_B = -c\Delta[1 + G(z)\{\Delta + [u(z) - v_B]\}]^{-1}, \quad (8.8.1)$$

and then neglect $[u(z) - v_B]$ in the denominator since it is $O(c^2)$ in the dilution limit ($c \rightarrow 0$). Then the $\mathbf{K} = 0$ element of (8.7.22) can be expanded as

$$\begin{aligned} \langle \langle \mathcal{R}(z) \rangle \rangle_{00} &= \{z - (E_0 + v_B) - [u(z) - v_B]\}^{-1} \\ &= (z - E_{B0})^{-1} - c\Delta(z - E_{B0})^{-2}[1 + \Delta G_B(z)]^{-1} + O(c^2), \end{aligned} \quad (8.8.2)$$

where

$$G_B(z) \equiv \mathcal{D}_B(z)_{nn} = N^{-1} \sum_{\mathbf{K}} (z - E_{B\mathbf{K}})^{-1}, \quad (8.8.3)$$

$$\mathcal{D}_B \equiv (z\mathbb{I} - \mathcal{H}_B)^{-1}, \quad (8.8.4)$$

$$E_{B\mathbf{K}} = E_B + t_{\mathbf{K}} = v_B + E_{\mathbf{K}} \quad (8.8.5)$$

are defined for *pure* B crystal. If there is a solution $E = E_b$, below the band edge E_{B0} of the B crystal, of the equation

$$1 - (-\Delta)G_B(E) = 0. \quad (E < E_{B0}), \quad (8.8.6)$$

the term linear in c in (8.8.2) contributes a line spectrum

$$F(E) = c[-dG_B(E_b)/dE_b]^{-1}(E_{B0} - E_b)^{-2}\delta(E - E_b). \quad (8.8.7)$$

The solution E_b represents a bound exciton state due to a guest atom A substituted for a host atom B, say at the site $n = 0$. In fact, the eigenequation for this system:

$$(\mathcal{H}_B + \mathbb{V})\psi = E\psi, \quad \mathbb{V}_{nm} = -\delta_{n0}\delta_{m0}\Delta \quad (8.8.8)$$

can be solved, with the use of (8.8.4), as

$$\psi_b(n) = -\Delta[\mathcal{D}_B(E_b)]_{n0}\psi_b(0). \quad (8.8.9)$$

The bound-state wave function $\psi_b(n)$ can be normalized with the use of the identity: $-[d\mathcal{D}_B(z)/dz]_{00} = \sum_n |\mathcal{D}_B(z)]_{n0}|^2$ which is obtained by differentiating (8.8.4). Noting that $\sum_n [\mathcal{D}_B(z)]_{n0} = [\mathcal{D}_B(z)]_{\mathbf{K}=0}$, one obtains

$$f_b \equiv \left| \sum_n \psi_b(n) \right|^2 = [-dG_B(E_b)/dE_b]^{-1}(E_{B0} - E_b)^{-2}. \quad (8.8.10)$$

The quantity f_b represents the oscillator strength of the optical transition to the bound exciton state in units of the atomic oscillator strength, as is obvious from comparison with (8.8.7), and also from the fact that

$$\left| \sum_n \psi_b(n) \right|^2 = N |\langle \psi_{\mathbf{K}=0} | \psi_b \rangle|^2. \quad (8.8.11)$$

The r.h.s. represents the oscillator strength borrowed from $\psi_{\mathbf{K}=0}$ into which the total oscillator strength of N atoms is concentrated in the pure crystal case. The l.h.s. of (8.8.11) represents the number M of those sites over which the bound-state

wave function extends with significant amplitude, since $\psi_b(n)$ is of the order of $M^{-1/2}$. The binding energy $\varepsilon_b \equiv E_{B0} - E_b$ and the orbital radius a_b of the bound exciton are related by $\varepsilon_b \propto a_b^{-2}$ due to the uncertainty principle. Hence one finds

$$f_b \sim M \propto (a_b)^d \propto (\varepsilon_b)^{-d/2}, \quad (8.8.12)$$

where d is the dimensionality of the crystal.

Thus, a shallowly bound exciton bears *giant oscillator strength*, corresponding to the number of atoms it covers. This fact was first pointed out by Rashba and Gurgenisvilli,²¹ and confirmed experimentally by Henry and Nassau²² through the measurement of the radiative lifetime of weakly bound excitons in CdS which was of the order of nanoseconds in agreement with theoretical estimation. The notion of giant oscillator strength plays another important role for the excitonic molecule, a pair of excitons weakly bound together, as will be mentioned in Chapter 15.

8.9 Amalgamation- and persistence-type excitons in mixed crystals

We present here the results of a model calculation for a Frenkel exciton in a mixed crystal using the method described in Section 8.7. The normalized density of states of the averaged crystal is assumed to be of the semi-spherical form

$$\rho^{(0)}(E) \equiv N^{-1} \sum_{\mathbf{K}} \delta(E - E_{\mathbf{K}}) = (2/\pi B^2) [B^2 - (E - \bar{E})^2]^{1/2}, \quad (8.9.1)$$

which rises at the bottom $(\bar{E} - B)$ and ends at the top $(\bar{E} + B)$ of the band in a way appropriate to the three-dimensional lattice (proportionally to the square-root energy). The state with $\mathbf{K} = 0$ is assumed to be at the band bottom: $t_{\mathbf{K}=0} = -B$. The function $G^{(0)}(z)$ defined by (8.7.17) is then given by

$$G^{(0)}(z) = \int dE \rho^{(0)}(E)/(z - E) = (2/B^2) \{ (z - \bar{E}) - [(z - \bar{E})^2 - B^2]^{1/2} \}, \quad (8.9.2)$$

where one is to take *that* square root with *negative* imaginary part.

Figure 8.7 shows how the density of states (dashed line), the absorption spectrum (solid line) and the imaginary part of the self-energy (dot-dashed line) of an exciton in a mixed crystal vary with $\Delta/2B$ for fixed $c = 0.5$. The A and B atomic states are *amalgamated* into a single band or *persist* as split bands according as $\Delta/2B$ is small or large. As shown by the dot-dashed lines, the perturbation is most remarkable near the (pseudo-) gap of the bands where the mixing of the two atomic states is greatest. The transition from the amalgamation (A-) type to the persistence (P-) type depends of course on the composition, as shown in Fig. 8.8. The amalgamated band and the split bands are shown against concentration c in Figs. 8.9 and 8.10. The lower edge of the amalgamated band, which is close to the absorption peak,

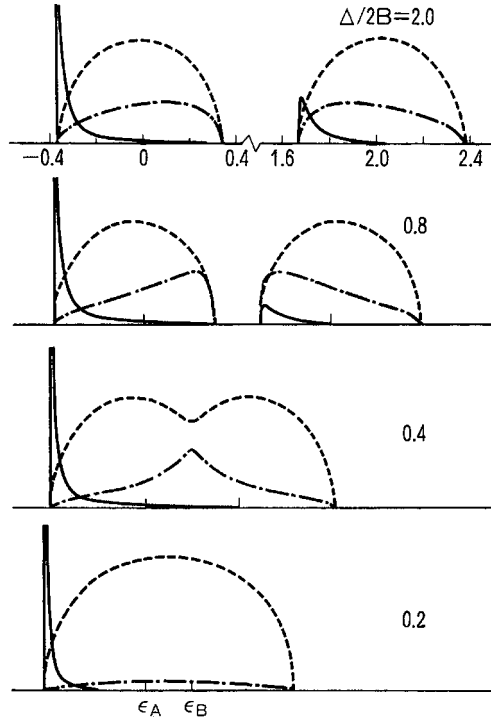


Fig. 8.7 Density of states (dashed lines), optical absorption spectrum (solid lines) and imaginary part of the self-energy (dot-dashed lines) of an exciton in 50:50 mixed crystals with various values of $\Delta/2B$.¹⁷

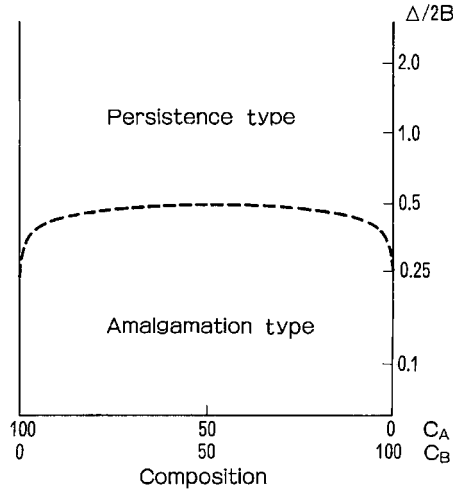


Fig. 8.8 Amalgamation and persistence regions in the coordinate space of composition and $\Delta/2B$.¹⁷

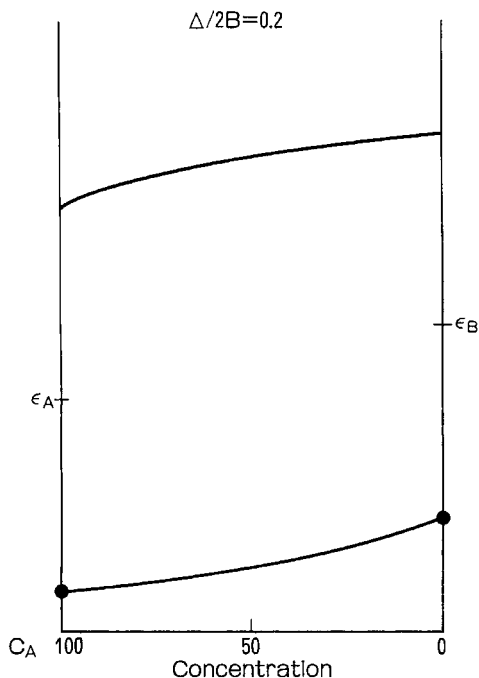


Fig. 8.9 Amalgamation band against concentration.¹⁷

shows downward quadratic bending against c from the linearly shifting averaged band edge, $E_{K=0} = c_A E_A + c_B E_B - B$, due to the self-energy shift given by

$$\text{Re}\{u(E_0)\} \sim c(1-c)\Delta^2 G^{(0)}(E_0), \quad [G^{(0)}(E_0) < 0]. \quad (8.9.3)$$

In Fig. 8.10 one sees how the impurity-bound excitons form the exciton band which finally grows into the host band as c increases. In this persistence case, the A and B absorption peaks which are close to the respective band bottoms shift to *opposite* directions and *nonlinearly*. In the intermediate case ($0.25 < \Delta/2B < 0.5$, see Fig. 8.8), the shallow bound exciton states form the impurity band which then merges into the host band, as shown in Fig. 8.11. One can clearly notice the giant oscillator strength characteristic of the weakly bound exciton; the ratio of the oscillator strengths of the impurity bands and the host band is much greater than the concentration ratio.

Before comparing the results of calculation with observations, we note that the present model applies also to the one-electron bands in mixed crystals as far as the density of states is concerned (except that there is no counterpart of the spectroscopically particular role of the exciton state $\mathbf{K} = 0$). Since the Wannier exciton is a composite particle comprising an electron in the conduction band and a positive hole in the valence band, one expects a correlation between the P–A alternative of

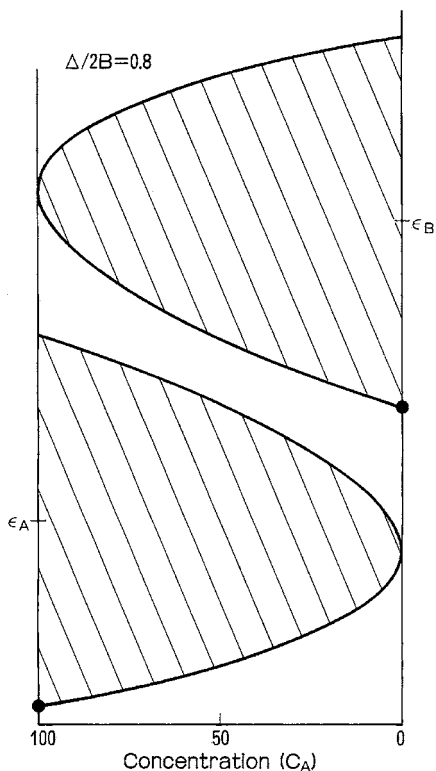


Fig. 8.10 Split band against concentration.¹⁷

the exciton and the combination of P–A alternatives of the valence and conduction bands. As a typical case, consider alkali halides (A.H.). Their valence bands consist mainly of halogen p-orbitals while their conduction bands consist mainly of alkali s-orbitals. Hence, the valence band is directly perturbed by halogen-mixing but not much by alkali-mixing, while the conduction band is directly perturbed by alkali-mixing but not much by halogen-mixing. Since the half-width B_v of the valence band is smaller than the electron affinity difference Δ_h between different halogen ions, the valence band in halogen-mixed A.H. belongs to P-type while the conduction band indirectly perturbed remains A-type. In contrast, the half-width B_c of the conduction band is larger than the ionization energy difference Δ_a between different alkali atoms, and remains A-type even against the direct perturbation by alkali-mixing. Namely, neither the valence nor the conduction band is brought into P-type by alkali-mixing, while only the valence band is brought into P-type by halogen-mixing. This provides a clear explanation for the experimental fact²³ that the exciton in alkali-mixed A.H. is A-type, but is only consistent with the fact²⁴ that the exciton in halogen-mixed A.H. is P-type since it is not immediately clear whether P–A combination of the two bands should result in a P-type exciton. A

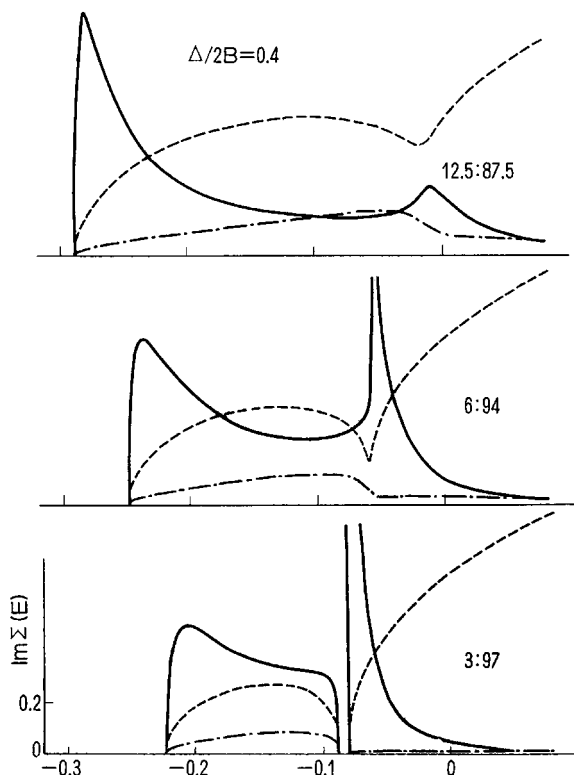


Fig. 8.11 Impurity band merging into host band with increase in concentration, $c_A : c_B$, for an intermediate value of $\Delta/2B = 0.4$. Energy (abscissa) is measured from E_A in units of $2B$. Line designations are the same as in Fig. 8.7.¹⁷

more plausible explanation of the latter fact is to consider the electron–hole relative motion in the Wannier exciton. Since the effective mass of the positive hole is much larger than that of the electron (corresponding to $B_v \ll B_c$ mentioned above), the orbit of the electron is much further out than that of the hole for a fixed position of the center of mass. This means that the hole is much more sensitive to the local potential than the electron which feels a potential averaged over many sites. To speak more generally, the Wannier exciton is governed by the perturbation on the band to which the heavier of the constituent particles belongs.

The downward bending of the peak position of the amalgamated exciton against the concentration c as predicted by Fig. 8.9 have been observed in a number of other ionic crystals and semiconductors. The $\text{AgBr}_{1-x}\text{Cl}_x$ ²⁵ and $\text{TlBr}_{1-x}\text{Cl}_x$ systems²⁶ have (see Chapter 10) indirect exciton absorption edges followed by direct absorption peaks, both of which are as a whole A-type with downward bending against x ; however, the indirect edges are complicated by the coexistent disorder-assisted no-phonon transitions and the phonon-assisted transitions (with phonon-related

A–P problems). Only a few examples of opposite shifts with x of the persistent peaks as predicted by Fig. 8.10 are known.²⁴

A systematic study of Wannier excitons in mixed crystals was made by Mahanti,²⁷ who analysed the linewidths of the spin–orbit components of $\text{CuCl}_{1-x}\text{Br}_x$ and emphasized the importance of single-particle (an electron or a hole) scattering by potential fluctuations and lattice vibrations.

Finally, it should be noted that the effect of impurity *clusters* are *not included* in the coherent potential approximation. In fact it is based upon the t-matrix (8.7.15) which describes exactly up to a single-impurity scattering but not more. Experimentally, however, the cluster effects seem to play important roles in the relatively low (but not too low) concentration region, especially in the P-type²⁸ or near P-type.²⁶

Extensive descriptions of the exciton spectra in pure and mixed molecular crystals are presented by Broude, Rashba and Sheka.²⁹

Polaron and the self-trapped state

9.1 Polarons in ionic crystals

The behavior of an electron in polarizable medium has a long history of study with versatile developments in condensed matter physics. In 1933 Landau¹ conceived of the *self-trapping* of an electron in a potential field of polarization induced by itself (*digging its own hole*, so to speak) as a possible origin of the color center in alkali halides (see Sections 4.7 and 4.8), recognized by their strong absorption bands in the visible region. Although the color center itself turned out later to be an electron at a lattice defect, his idea was further developed by Pekar,² Fröhlich,³ Lee and Pines,⁴ Feynman⁵ and others,^{6,7} partly stimulated by the field-theoretical approaches to an electron in a self-induced electromagnetic field in a vacuum, and was applied to similar or related problems in condensed matter physics. We shall describe some of these developments and applications in this chapter, with the use of the dielectric theory developed in Chapters 5 and 6.

Let us consider the fluctuating electric field in an isotropic insulator caused by lattice vibrations. The fluctuation–dissipation theorem allows us to express it in terms of the dielectric dispersion as follows.⁸ In the absence of an external charge, eq. (1.1.3) with $\mathbf{D} = \epsilon_e \mathbf{E} + \mathbf{P}$ gives $\epsilon_e \mathbf{k} \cdot \mathbf{E}_k = -\mathbf{k} \cdot \mathbf{P}_k$. The polarization density \mathbf{P} due to lattice vibrations causes electrostatic potential $\phi(\mathbf{r})$ through $\mathbf{E} = -\nabla \phi(\mathbf{r})$. The latter can be written, in the Fourier expansion in conformity with (6.1.27), as

$$\phi(\mathbf{r}) = V^{-1/2} \sum_{\mathbf{k}} \phi_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r}) = \epsilon_e^{-1} V^{-1/2} \sum_{\mathbf{k}} (-i/k) \mathbf{k} \cdot \mathbf{P}_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r}). \quad (9.1.1)$$

Only the longitudinal part of the Fourier component $\mathbf{P}_{\mathbf{k}}$ given by (6.4.2), and hence $q_{\parallel}^{(s)}(\mathbf{k}) \equiv \mathbf{k} \cdot \mathbf{q}^{(s)}(\mathbf{k})$, is needed in (9.1.1). The latter can be obtained as follows. Corresponding to (5.3.7) to be obtained by the \mathbf{E} method (6.3.2), the \mathbf{D} method (6.3.3) gives

$$1/\epsilon_e - 1/\epsilon(\mathbf{k}, \omega) = \epsilon_e^{-2} \sum_s N_0 [q_{\parallel}^{(s)}(\mathbf{k})]^2 / (\omega_s^2 - \omega^2). \quad (9.1.2)$$

Noting that the s th term becomes dominant at $\omega \sim \omega_s$, one obtains

$$[\partial_\epsilon(\mathbf{k}, \omega)/\partial(\omega^2)]_{\omega=\omega_s} = \epsilon_e^2/N_0[q_\parallel^{(s)}]^2. \quad (9.1.3)$$

Putting (6.4.2) and (9.1.3) into (9.1.1), one obtains

$$\begin{aligned} \phi(\mathbf{r}) = & V^{-1/2} \sum_{\mathbf{k}s} [\hbar/\partial\epsilon(\mathbf{k}, \omega_s)/\partial\omega_s]^{1/2} \\ & \times (-i/k)[b_{s,\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r}) - b_{s,\mathbf{k}}^\dagger \exp(-i\mathbf{k} \cdot \mathbf{r})] \end{aligned} \quad (9.1.4)$$

where the sign of $q_\parallel^{(s)}$ makes no difference since it can be absorbed in the definition of $b_{s,\mathbf{k}}$. The transverse modes with $\omega_s = \Omega_s$, for which $\partial\epsilon/\partial\omega_s = \infty$ (see Fig. 5.4(a)), do not contribute to the electrostatic potential as is naturally expected, so that one can confine oneself to longitudinal modes which will be numbered ℓ . Denoting the wave vector of a phonon by \mathbf{q} to distinguish it from that of an electron, \mathbf{k} , one can write the interaction of the electron with optical phonons as

$$\begin{aligned} H_I = (-e)\phi(\mathbf{r}) = & (-e)V^{-1/2} \sum_{\ell\mathbf{q}} [\hbar/\partial\epsilon(\mathbf{q}, \omega_\ell)/\partial\omega_\ell]^{1/2} \\ & \times (-i/q)[b_{\ell\mathbf{q}} \exp(i\mathbf{q} \cdot \mathbf{r}) - b_{\ell\mathbf{q}}^\dagger \exp(-i\mathbf{q} \cdot \mathbf{r})]. \end{aligned} \quad (9.1.5)$$

Due to the factor q^{-1} (which originates from the Fourier component of the dipolar potential $\propto r^{-2}$ of the polarization due to optical lattice vibrations), only those phonons with small q have significant interaction (*long-range* interaction) with the electron. One can therefore put $q \cong 0$ in $\epsilon(\mathbf{q}, \omega_\ell)$ and rewrite it as $\epsilon(\omega_\ell)$.

In the case of one longitudinal optical mode, one can omit summation over ℓ and rewrite (9.1.5), with the use of (5.3.14) to (5.3.16), as

$$\begin{aligned} H_I = (-e)\phi(\mathbf{r}) = & [(\epsilon_e^{-1} - \epsilon_s^{-1})e^2\hbar\omega_\ell/2]^{1/2} \\ & \times V^{-1/2} \sum_{\mathbf{q}} (i/q)[b_{\mathbf{q}} \exp(i\mathbf{q} \cdot \mathbf{r}) - b_{\mathbf{q}}^\dagger \exp(-i\mathbf{q} \cdot \mathbf{r})]. \end{aligned} \quad (9.1.6)$$

The Hamiltonian which has been used in the polaron with the effective mass approximation described in Section 7.4 is as follows:

$$H = h_c + H_L + H_I, \quad h_c = \mathbf{p}^2/2m_e, \quad H_L = \hbar\omega_\ell \sum_{\mathbf{q}} b_{\mathbf{q}}^\dagger b_{\mathbf{q}}, \quad (9.1.7)$$

where $\mathbf{p} \equiv -i\hbar\partial/\partial\mathbf{r} = -i\hbar\nabla$ is the *pseudo-momentum* of the electron in the conduction band, and $m_e \equiv m_c$ is its *effective mass* at the band bottom from which one-electron energy h_c is to be measured (see eqs. (7.4.9, 7.4.10)). The total *pseudo-momentum* defined by

$$\mathbf{P} = \mathbf{p} + \sum_{\mathbf{q}} \hbar\mathbf{q} b_{\mathbf{q}}^\dagger b_{\mathbf{q}} \quad (9.1.8)$$

turns out to be a constant of motion, namely a quantity to be conserved throughout the motion, since it commutes with H :

$$[\mathbf{P}, H]_- \equiv [\mathbf{P}, h_c + H_L + H_I]_- = [\mathbf{P}, H_I]_- = 0 \quad (9.1.9)$$

as is readily confirmed by operator algebra, or from the statement in the following paragraph. Within a crystal, the pseudo-momentum, which is defined as \hbar times the wave vector \mathbf{k} or \mathbf{q} , must be distinguished from the true momentum. A more exact statement is that the sum of the wave vectors of an electron and phonons is conserved. This statement is certainly correct with the Hamiltonian (9.1.6, 9.1.7), but more generally invalidated due the Umklapp processes. We will hereafter mostly omit the term “pseudo”, confining ourselves to the former situation. As for the distinction between the pseudo-momentum and the true momentum, a more elaborate argument is given by Fröhlich.³

The eigenstate of H_L with n_q phonons with wave vector \mathbf{q} , $n_{q'}$ phonons with wave vector \mathbf{q}' and so on, can be expressed as

$$|\dots, n_q, \dots\rangle = \prod_q (n_q!)^{-1/2} (b_q^\dagger)^{n_q} |\dots, 0, \dots\rangle \quad (9.1.10)$$

where the product is to run over all \mathbf{q} s within the first Brillouin zone. The no-phonon state $|\dots, 0, \dots\rangle$, also called the “phonon vacuum state”, will hereafter be abbreviated as $|0\rangle$. The normalized wave function of the electron in the eigenstate $|\mathbf{k}\rangle$ of H_e with wave vector \mathbf{k} (and pseudo-momentum $\hbar\mathbf{k}$) is given by $\langle \mathbf{r} | \mathbf{k} \rangle = V^{-1/2} \exp(i\mathbf{k} \cdot \mathbf{r})$. The direct product of this electron state with the phonon state mentioned above can then be expressed as $|\mathbf{k}; \dots, n_q, \dots\rangle \equiv |\mathbf{k}\rangle |\dots, n_q, \dots\rangle$. The first and second terms in $[\dots]$ of the second line of (9.1.6) have the effect of scattering this electron to states with wave vector $\mathbf{k}' = \mathbf{k} \pm \mathbf{q}$ by absorbing or emitting a phonon with wave vector \mathbf{q} , respectively, through the transition matrix elements:

$$\begin{aligned} & \langle \mathbf{k} \pm \mathbf{q}; \dots, n_q + 1, \dots | H_I | \mathbf{k}; \dots, n_q, \dots \rangle \quad (n_{q'} \text{ with other } \mathbf{q}' \text{ unchanged}) \\ &= [(\epsilon_e^{-1} - \epsilon_s^{-1})e^2\hbar\omega_\ell/2]^{1/2} V^{-1/2} (\pm i/q)(n_q \text{ or } n_q + 1)^{1/2}, \end{aligned} \quad (9.1.11)$$

where the value of (9.1.8) is unchanged (the same between the states before and after the transition). The value is kept constant for any times of scattering by phonons with different \mathbf{q} s.

The rate of scattering is given, by Fermi's golden rule for the first-order transition, as

$$\begin{aligned} (1/\tau)_{\text{abs. or em.}} &= (2\pi/\hbar) [(\epsilon_e^{-1} - \epsilon_s^{-1})e^2\hbar\omega_\ell/2] V^{-1} \sum_q q^{-2} \\ &\times (n_q \text{ or } n_q + 1) \delta[(\hbar^2/2m_e)\{(\mathbf{k} \pm \mathbf{q})^2 - \mathbf{k}^2\} \mp \hbar\omega_\ell]. \end{aligned} \quad (9.1.12)$$

The thermal averages of the phonon number and the electron kinetic energy in the initial state are given respectively by

$$\langle\langle n_q \rangle\rangle = [\exp(\beta\hbar\omega_\ell) - 1]^{-1} \equiv \langle\langle n \rangle\rangle, \quad \langle\langle (\hbar\mathbf{k})^2/2m_e \rangle\rangle = (3/2)\beta^{-1}. \quad (9.1.13)$$

At low temperatures such that $\beta\hbar\omega_\ell \gg 1$, only the scattering by phonon *absorption* can take place due to the energy conservation, and the rate is given by

$$(1/\tau) \simeq (1/\tau)_{\text{abs.}} = 2\alpha\omega_\ell\langle\langle n \rangle\rangle, \quad (9.1.14)$$

where α , the polaron coupling constant, is given by

$$\alpha \equiv (1/2)[(e^2k_p/4\pi)(\epsilon_e^{-1} - \epsilon_s^{-1})]/\hbar\omega_\ell, \quad (9.1.15)$$

$$\text{with } k_p \equiv (2m_e\omega_\ell/\hbar)^{1/2}. \quad (9.1.16)$$

By phonon absorption, the electron obtains enough kinetic energy to emit a phonon. Since the rate of emission is of the order of $\alpha\omega_\ell(\langle\langle n \rangle\rangle + 1)$ and much larger than (9.1.14), the phonon absorption is so immediately followed by the phonon emission that (9.1.14) can be taken to give the rate of the resonant two-phonon scattering process. Since the electron in the final state of this scattering has almost isotropic distribution of momentum (the statistical average of $\cos\theta$ of the scattering angle θ is nearly zero), the mobility μ_c (defined as the mean velocity of an electron divided by the applied electric field) of the electron is given by $e\tau/m_e = (e/2m_e\alpha\omega_\ell)[\exp(\beta\hbar\omega_\ell) - 1]$. This temperature dependence of the electron mobility has in fact been confirmed in many ionic crystals and, combined with the determination of the effective mass m_c by cyclotron resonance, served to estimate the value of coupling constant α . However, the effective mass as well as the mobility calculated above by perturbation theory must be corrected due to the large value of α (\sim or > 1) in ionic crystals due to the polaron effect to be studied here.

As will be seen below, k_p is the reciprocal radius of the polaron, the phonon-dressed electron. The dimensionless quantity α defined by (9.1.15) is the electron-optical phonon coupling constant which is most popularly used in the polaron theory. The interaction Hamiltonian (9.1.6) can then be rewritten as

$$H_1 = (-e)\phi(\mathbf{r}) = \sum_{\mathbf{q}} V_{\mathbf{q}} b_{\mathbf{q}} \exp(i\mathbf{q} \cdot \mathbf{r}) + V_{\mathbf{q}}^* b_{\mathbf{q}}^\dagger \exp(-i\mathbf{q} \cdot \mathbf{r}), \quad (9.1.17)$$

$$V_{\mathbf{q}} \equiv i[4\pi\alpha]^{1/2}\hbar\omega_\ell V^{-1/2}k_p^{-1/2}q^{-1}. \quad (9.1.18)$$

Let us now study more elaborately the state of an electron in the phonon field within the perturbation theory, confining ourselves to absolute zero of temperature where there are no *real* phonons. However, the electron induces *virtual* phonons in the following sense. Starting from the *bare* electron state $|\mathbf{k}; \dots, 0_{\mathbf{q}}, \dots\rangle$, one

finds the wave function for the electron–phonon system, correct up to the first-order perturbation, as

$$|\Psi_k\rangle = |\mathbf{k}; \dots, 0_q, \dots\rangle + \sum_q [(\hbar^2/2m_e)\{\mathbf{k}^2 - (\mathbf{k} - \mathbf{q})^2\} - \hbar\omega_\ell]^{-1} \times V_q^* |\mathbf{k} - \mathbf{q}; \dots, 1_q, \dots\rangle. \quad (9.1.19)$$

Namely, the states with one phonon become mixed into the bare electron state due to the electron–phonon interaction H_I . The electron is, so to speak, dressed with phonons which are *virtually* emitted (that is, without satisfying the energy conservation). Such a dressed particle was named a “polaron”, implying that quanta of electric *polarization* are induced around the *electron*. What are the effects of such dressing?

In the first place, the electron with wave vector \mathbf{k} is subject to an energy shift Δ_k , the second-order term of the perturbation theory (note that the first-order term vanishes because H_I has no diagonal element):

$$\Delta_k = \sum_q |V_q|^2 / [(\hbar^2/2m_e)\{\mathbf{k}^2 - (\mathbf{k} - \mathbf{q})^2\} - \hbar\omega_\ell]^{-1}. \quad (9.1.20)$$

After integration on the direction of \mathbf{q} , one obtains

$$\begin{aligned} \Delta_k &= -\alpha\hbar\omega_\ell(2\pi)^{-1} \int_0^\infty dq (k_p/2kq) \ln \{[(k_p^2 + q^2) + 2kq]/[(k_p^2 + q^2) - 2kq]\} \\ &= -\alpha\hbar\omega_\ell \sum_{\nu=0}^\infty (2\nu + 1)^{-1} (2\nu)! [(2)^{2\nu} (\nu!)^2]^{-1} (k/k_p)^{2\nu}, \end{aligned} \quad (9.1.21)$$

where the radius of convergence is readily shown to be $(k/k_p) = 1$. For $k \ll k_p$ one can expand

$$\Delta_k = -\alpha\hbar\omega_\ell - (\alpha/6)(\hbar^2 k^2/2m_e). \quad (9.1.22)$$

Adding this second-order correction energy to the bare electron energy, one finds that

$$E_k = -\alpha\hbar\omega_\ell + \hbar^2 k^2/2m_p, \quad (9.1.23)$$

$$\text{where } m_p = m_e/(1 - \alpha/6). \quad (9.1.24)$$

Corresponding to the mass correction (9.1.24), the threshold wave number beyond which the electron can emit an optical phonon, as was initially given by k_p of (9.1.16), should be corrected to a greater value.

The expectation value n of the total number of virtual phonons in the polaron state (9.1.19) is similarly calculated as

$$\begin{aligned} n &\equiv \langle \Psi_k | \sum_q b_q^\dagger b_q | \Psi_k \rangle = \sum_q |V_q|^2 / [(\hbar^2/2m_e)\{\mathbf{k}^2 - (\mathbf{k} - \mathbf{q})^2\} - \hbar\omega_\ell]^{-2} \\ &= (\alpha/2)[1 - (k/k_p)^2]^{-1/2}. \end{aligned} \quad (9.1.25)$$

Let us rewrite (9.1.19), by retaining the electron coordinate \mathbf{r} , as

$$\Psi_{\mathbf{k}}(\mathbf{r}) = V^{-1/2} \exp(i\mathbf{k} \cdot \mathbf{r}) \left[1 + \sum_{\mathbf{q}} \{(\hbar^2/2m_e)[\mathbf{k}^2 - (\mathbf{k} - \mathbf{q})^2] - \hbar\omega_{\ell}\}^{-1} \right. \\ \left. \times V_{\mathbf{q}}^* \exp(-i\mathbf{q} \cdot \mathbf{r}) b_{\mathbf{q}}^{\dagger} \right] |0\rangle \quad (9.1.26)$$

with the use of the phonon vacuum state $|0\rangle$, and consider the electrostatic potential $\phi(\mathbf{r})$ (see eq. (9.1.6) of the polarization induced by the electron momentarily at \mathbf{r}_e . Then the quantity $\langle \Psi_{\mathbf{k}}(\mathbf{r}_e) | \phi(\mathbf{r}) | \Psi_{\mathbf{k}}(\mathbf{r}_e) \rangle$, which is a function of the relative coordinate $\mathbf{r} - \mathbf{r}_e$ only, represents the self-induced potential, at \mathbf{r} , of the electron at \mathbf{r}_e , (except for the normalization constant which can be taken to be unity in the perturbation limit). Keeping $(\mathbf{r} - \mathbf{r}_e) \equiv \boldsymbol{\rho}$ constant and performing the remaining integration in \mathbf{r}_e over the volume V , one obtains the self-induced potential:

$$U(\mathbf{r} - \mathbf{r}_e) = \int_v d\mathbf{r}_e \langle \Psi_{\mathbf{k}}(\mathbf{r}_e) | \phi(\mathbf{r}) | \Psi_{\mathbf{k}}(\mathbf{r}_e) \rangle \\ = -(-e)^{-1} \sum_{\mathbf{q}} |V_{\mathbf{q}}|^2 2 \cos(\mathbf{q} \cdot \boldsymbol{\rho}) [(\hbar^2/2m_e)\{\mathbf{k}^2 - (\mathbf{k} - \mathbf{q})^2\} - \hbar\omega_{\ell}]^{-1} \\ \simeq (4/\pi)(\alpha/e) \hbar\omega_{\ell} k_p \int_0^{\infty} dq [\sin(q\rho)/q\rho] / [k_p^2 + q^2]^{-1} \quad (k \ll k_p) \\ = -(-e/4\pi)(\epsilon_e^{-1} - \epsilon_s^{-1})[1 - \exp(-k_p\rho)]/\rho. \quad (9.1.27)$$

One finds that $U(\boldsymbol{\rho})$ tends to $-(-e/4\pi)(\epsilon_e^{-1} - \epsilon_s^{-1})/\rho$ at long distance as is expected for the self-induced potential by a *point* charge, but that it deviates therefrom at short distance, tending to a finite value corresponding to the replacement of ρ by k_p^{-1} . Therefore, k_p^{-1} represents, so to speak, the radius of the polaron which is governed by the uncertainty principle for the electron with finite mass.

It is instructive to recapitulate here Fröhlich's physical argument on the polaron radius.³ Let us consider a dynamical problem of an electron with velocity v interacting with the lattice with natural frequency ω_{ℓ} . The electron travels a distance v/ω_{ℓ} within one oscillation of the lattice. As seen at a distance d_1 , the electron looks like a *static* charge if the angular distance $v/\omega_{\ell}d_1$ of the electron movement is much smaller than unity. On the other hand, the electron has spatial extension of at least $d_2 \sim \hbar/m_e v$ due to the uncertainty principle. The *minimum* distance d which satisfies both requirements, namely, which enables one to see the charge as static from its outside, is given by $d \sim k_p^{-1}$ corresponding to the velocity $v \sim \hbar k_p/m_e$, in agreement with the above argument.

As is obvious from the definition of k_p , (9.1.16), $\hbar k_p$ represents the recoil momentum of a slow electron due to its virtual emission of a phonon. The electron-phonon interaction, whose intensity is proportional to $|V_{\mathbf{q}}|^2 \propto q^{-2}$, is further enfeebled by the factor $(k_p/q)^2$ (a kind of motional narrowing which will be studied

in Chapter 10) when the recoil momentum $\hbar q$ is much greater than $\hbar k_p$, as can be seen from the energy denominators in eqs. (9.1.20). The effective range of coupling is thus confined within a small region of the first Brillouin zone since k_p is several to ten times smaller than the reciprocal lattice k_0 in typical insulators. This justifies the use of the electron–phonon interaction (9.1.6) which was derived in the long-wavelength limit with the dielectric continuum model ignoring the discrete lattice structure as well as the dispersion of $\omega_{\ell k}$.

It is to be noted that the negative energy shift given by (9.1.21) and the total number of virtual phonons given by (9.1.25), both being the results of the lowest-order perturbation theory, diverge to ∞ as k tends to k_p . While the divergence itself is obviously an artifact of perturbation theory, this breakdown is related with the fact that the electron with $k > k_p$ can emit a phonon as a real process with finite lifetime τ even at absolute zero of temperature. Namely, the electron is no longer in a stationary state with a well-defined energy but in a decaying state with finite level width \hbar/τ .

Confining ourselves to the slow polaron with $k \ll k_p$, we must note that it is not the bare mass m_c but the renormalized mass m_p of the polaron which is experimentally observed, say, by cyclotron resonance. One may well wonder if the enhanced mass, given by (9.1.24) within the lowest-order perturbation theory, tends to diverge for a large coupling constant α . This is not necessarily an unrealistic question since α may be of the order of unity in typical ionic crystals, and can be as large as several to tens in narrow bands with m_c much larger than the free-electron mass. For this reason, it is very important to improve or extend the perturbation theory so as to cover the case of a larger coupling constant which was in fact the primary motivation for this study.

9.2 Polaron theories for intermediate to strong coupling ranges

The intermediate coupling theory presented independently by Lee and Pines,⁴ Gurari⁹ and Tiablikov¹⁰ was a useful step towards extending the validity range of the polaron theory beyond the perturbation theory. Being a variational theory which gives an upper bound of energy, it can avoid the dangerous divergences inherent to the perturbation theory, providing us with more reliable judgements on what really is going on. We will give here the main point of it briefly.

Since the total momentum \mathbf{P} given by (9.1.8) commutes with Hamiltonian H (see (9.1.9)), one can diagonalize them simultaneously. This can be performed as follows. Let us rewrite the Schrödinger equation

$$H\Psi = E\Psi \quad (9.2.1)$$

which we have been studying, as a canonical transformation generated by

$$U_1 = \exp \left[-i \sum_q (\mathbf{q} b_q^\dagger b_q) \cdot \mathbf{r} \right] \quad (U_1^\dagger = U_1^{-1}). \quad (9.2.2)$$

The new wave function Φ and the new Hamiltonian \bar{H} defined by

$$\Phi \equiv U_1^{-1} \Psi, \quad \bar{H} \equiv U_1^{-1} H U_1 \quad (9.2.3)$$

must then satisfy the Schrödinger equation

$$\bar{H} \Phi = E \Phi. \quad (9.2.4)$$

The individual variables which appeared in (9.1.6, 9.1.7) are transformed as $\bar{\mathbf{p}} \equiv U_1^{-1} \mathbf{p} U_1 = \mathbf{p} - \sum_q \hbar \mathbf{q} b_q^\dagger b_q$, $\bar{b}_q \equiv U_1^{-1} b_q U_1 = b_q \exp(-i \mathbf{q} \cdot \mathbf{r})$, $\bar{\mathbf{P}} \equiv U_1^{-1} \mathbf{P} U_1 = \mathbf{P}$, etc., as can be derived from the commutation relation: $[\mathbf{p}, U_1]_- = -i \hbar \partial U_1 / \partial \mathbf{r}$ and from the relation $b f(n) = f(n+1) b$ which is a generalization of the commutation relation $b n = (n+1) b$ ($n \equiv b^\dagger b$) to an arbitrary function $f(n)$, whereas $\exp(-i \mathbf{q} \cdot \mathbf{r})$ is unchanged because of its commutativity with U_1 . \bar{H} can therefore be written as

$$\bar{H} = \sum_q \hbar \omega_{\ell} b_q^\dagger b_q + \sum_q (V_q b_q + V_q^* b_q^\dagger) + \left(\mathbf{p} - \sum_q \hbar \mathbf{q} b_q^\dagger b_q \right)^2 / 2m_e. \quad (9.2.5)$$

Since the electron coordinate \mathbf{r} is already eliminated in (9.2.5), the only variable \mathbf{p} which does not commute with \mathbf{r} can be taken as a c-number. Because of the above relations, this \mathbf{p} is nothing other than the total momentum \mathbf{P} . The expression (9.2.5) with \mathbf{p} replaced by \mathbf{P} can therefore be taken as the Hamiltonian for the phonon state in the subspace of each eigenvalue of \mathbf{P} .

The next step is to solve the eigenvalue problem for \bar{H} . Let us put $\Phi = U_2 |0\rangle$ where $|0\rangle$ represents the phonon vacuum state and U_2 is a unitary operator which is a function of the phonon variables b_q and b_q^\dagger . As an approximation, one takes a variational method with a trial form:

$$U_2 = \exp \left[\sum_q \{ b_q^\dagger f(\mathbf{q}) - b_q f^*(\mathbf{q}) \} \right], \quad (9.2.6)$$

and chooses a trial function $f(\mathbf{q})$ so as to minimize the energy

$$E = \langle 0 | U_2^{-1} \bar{H} U_2 | 0 \rangle. \quad (9.2.7)$$

Transformation (9.2.6) has the effect of *displacing* the harmonic oscillators:

$$U_2^{-1} b_q U_2 = b_q + f(\mathbf{q}), \quad U_2^{-1} b_q^\dagger U_2 = b_q^\dagger + f^*(\mathbf{q}) \quad (9.2.8)$$

as is obvious from the relation $[b_q, U_2]_- = (\partial / \partial b_q^\dagger) U_2$ which is a generalization of the commutation relation $[b_q, b_q^\dagger b_q]_- = b_q = (\partial / \partial b_q^\dagger) b_q^\dagger b_q$ to an arbitrary function U_2 . The transformed Hamiltonian $\bar{\bar{H}} \equiv U_2^{-1} \bar{H} U_2$, which can be readily calculated,

consists of terms up to the fourth power in the bs . Decomposing it to $\overline{\overline{H}}_1$ which contributes to (9.2.7) and the remainder $\overline{\overline{H}}_2$, one obtains

$$\begin{aligned} E = \langle 0 | \overline{\overline{H}}_1 | 0 \rangle &= P^2/2m_e + \sum_q [V_q f(\mathbf{q}) + V_q^* f^*(\mathbf{q})] \\ &+ (\hbar^2/2m_e) \left[\sum_q |f(\mathbf{q})|^2 \right]^2 \\ &+ \sum_q |f(\mathbf{q})|^2 [\hbar\omega_\ell - \hbar\mathbf{q} \cdot \mathbf{P}/m_e + \hbar^2 q^2/2m_e]. \end{aligned} \quad (9.2.9)$$

In order to minimize (9.2.9), one has to solve $\delta E/\delta f(\mathbf{q}) = \delta E/\delta f^*(\mathbf{q}) = 0$, namely

$$\begin{aligned} V_q + f^*(\mathbf{q}) \left[\hbar\omega_\ell - \hbar\mathbf{q} \cdot \mathbf{P}_{\text{tot}}/m_e + \hbar^2 q^2/2m_e \right. \\ \left. + (\hbar^2/m_e) \mathbf{q} \cdot \left\{ \sum_{q'} |f(\mathbf{q}')|^2 \mathbf{q}' \right\} \right] = 0. \end{aligned} \quad (9.2.10)$$

Since \mathbf{P} is the only direction specified in the above expression, one can put

$$\sum_q |f(\mathbf{q})|^2 \hbar\mathbf{q} = \eta \mathbf{P}. \quad (9.2.11)$$

Putting (9.2.10) in (9.2.11), one obtains the equation for η :

$$(\eta - 1)^2 \eta = (\alpha/2)(P/\hbar k_p)^{-3} [\sin^{-1} u - u(1 - u^2)^{-1/2}], \quad (9.2.12)$$

$$u \equiv (\eta - 1)P/\hbar k_p. \quad (9.2.13)$$

The energy, given by

$$E = (P^2/2m_e)(1 - \eta^2) - \alpha \hbar\omega_\ell (\sin^{-1} u)/u, \quad (9.2.14)$$

can be expanded, when $P \ll \hbar k_p$, as

$$E = -\alpha \hbar\omega_\ell + P^2/[2m_e(1 + \alpha/6)] + O(P/\hbar k_p)^4 \hbar\omega_\ell, \quad (9.2.15)$$

whereas the number of virtual phonons is given by

$$\begin{aligned} \langle \langle n \rangle \rangle &= \langle \Phi | \sum_q b_q^\dagger b_q | \Phi \rangle = \sum_q |f(\mathbf{q})|^2 \\ &= (\alpha/2)[1 + (1/2)(P/\hbar k_p)^2(1 + \alpha/6)^{-2}] + O(P/\hbar k_p)^4. \end{aligned} \quad (9.2.16)$$

Although the P -independent term of energy (9.2.15) obtained by the variational method is the same as that obtained by the perturbation theory, it is important that (9.2.15) gives an upper bound for energy. In view of the fact that the variational calculation with up to one-phonon states considered gives energy $E = -x\hbar\omega_\ell > -\alpha\hbar\omega_\ell$ where x is the positive root of the equation $x(1 + x)^{1/2} = \alpha$, the states with two and more phonons automatically included in the variational wave function $\Phi = U_2|0\rangle$ contribute a significant part of the self-energy $-\alpha\hbar\omega_\ell$ even at $\alpha \sim 2$. The effective mass, $m_p = m_e(1 + \alpha/6)$ of (9.2.15), is smaller and less divergent than $m_p = m_e(1 - \alpha/6)^{-1}$ predicted by the perturbation theory, although they are equal up to the first-order term in α , as they should be.

Although we did not give an expression for $\overline{\overline{H}}_1$, it can be shown that it is diagonalized together with $b_q^\dagger b_q$ when (9.2.10) is satisfied and, hence, the trial function $\Phi = U_2|0\rangle$ is an exact eigenfunction of $\overline{\overline{H}}_1$ under the same condition. Therefore, the second-order perturbation energy due to $\overline{\overline{H}}_2$ (the first-order energy vanishes) will provide a judgement on the accuracy of the solution obtained above. In fact, such an estimation was made by Lee, Low and Pines,¹¹ to give

$$\Delta E = -0.01592\alpha^2\hbar\omega_\ell + 0.02\alpha^2(1 + \alpha/6)^{-2}P^2/2m_e, \quad (9.2.17)$$

where the first term gives an exact coefficient of the α^2 term in the conventional perturbation theory.

The approach from the strong coupling limit was started by Pekar,² following Landau's picture that an electron is captured by the electrostatic potential induced by itself. He assumed a product form: $\Psi = \psi\Phi$, of the electronic wave function $\psi(\mathbf{r})$ and the phonon wave function Φ , which are to be chosen so as to minimize the expectation value of the total energy (9.1.7) under the normalization conditions: $\langle\psi|\psi\rangle_e = \int d\mathbf{r}|\psi(\mathbf{r})|^2 = 1$, and $\langle\Phi|\Phi\rangle_L = 1$. The resulting simultaneous equations can be written, with the use of Lagrange's undetermined multipliers λ_1 and λ_2 , as

$$[-(\hbar^2/2m_e)\nabla^2 + \langle\Phi|H_I|\Phi\rangle_L]\psi = \lambda_1\psi, \quad (9.2.18)$$

$$\left[\sum_q \hbar\omega_\ell b_q^\dagger b_q + \int \psi H_I \psi d\mathbf{r}\right]\Phi = \lambda_2\Phi. \quad (9.2.19)$$

The former represents the Schrödinger equation for the electron in the electrostatic field of the displaced lattice while the displacements are determined by the latter equation. In fact the integral is given by

$$\int \psi H_I \psi d\mathbf{r} = \sum_q V_q b_q \int d\mathbf{r} |\psi(\mathbf{r})|^2 \exp(i\mathbf{q} \cdot \mathbf{r}) + (\text{h.c.})$$

and hence $[\dots]$ on the l.h.s. of (9.2.19) can be written as

$$\sum_q \hbar\omega_\ell b_q^\dagger b_q' - \sum_q \hbar\omega_\ell |\langle b_q \rangle|^2, \quad (9.2.20)$$

$$b_q' \equiv b_q - \langle b_q \rangle, \quad (9.2.21)$$

$$\langle b_q^\dagger \rangle = \langle b_q \rangle^* = -(\hbar\omega_\ell)^{-1} V_q \int d\mathbf{r} |\psi(\mathbf{r})|^2 \exp(i\mathbf{q} \cdot \mathbf{r}). \quad (9.2.22)$$

In the lowest vibrational state $\Phi'_0(n_q' \equiv b_q^\dagger b_q' = 0)$ of the displaced harmonic oscillators in (9.2.20), the self-consistent field in (9.2.18) is given by

$$\langle\Phi'_0|H_I|\Phi'_0\rangle_L = (\epsilon_\infty^{-1} - \epsilon_s^{-1}) \int d\mathbf{r}' e^2 |\psi(\mathbf{r}')|^2 / 4\pi |\mathbf{r} - \mathbf{r}'|. \quad (9.2.23)$$

The expectation value of the total energy is then given by

$$\begin{aligned}
 E &= \langle \psi | \Phi_0' | H | \psi | \Phi_0' \rangle = -(\hbar^2/2m_e) \int d\mathbf{r} \psi^* \nabla^2 \psi - \sum_{\mathbf{q}} \hbar \omega_{\ell} |\langle b_{\mathbf{q}} \rangle|^2 \\
 &= -(\hbar^2/2m_e) \int d\mathbf{r} \psi^* \nabla^2 \psi - (\epsilon_{\infty}^{-1} - \epsilon_s^{-1}) \\
 &\quad \times \iint d\mathbf{r} d\mathbf{r}' e^2 |\psi(\mathbf{r})|^2 |\psi(\mathbf{r}')|^2 / 4\pi |\mathbf{r} - \mathbf{r}'|.
 \end{aligned} \tag{9.2.24}$$

Solving the self-consistent non-linear Schrödinger equation (9.2.18) with (9.2.23) is equivalent to minimizing the total energy (9.2.24). If one puts a trial wave function of the Gaussian form:

$$\psi(\mathbf{r}) = (\mu^2/\pi)^{3/4} \exp(-\mu^2 r^2/2), \tag{9.2.25}$$

into (9.2.24) and minimizes with respect to μ , one obtains

$$E = -(3\pi)^{-1} \alpha^2 \hbar \omega_{\ell}, \tag{9.2.26}$$

which gives lower, and hence, better energy than the corresponding value $-\alpha \hbar \omega_{\ell}$ (see (9.2.15)) of the intermediate coupling theory at $\alpha > 9.42$. The expectation value of the radius of the orbital (9.2.25) is given by

$$\langle r \rangle = 2\pi^{-1/2} \mu^{-1} = 4.24(\alpha k_p)^{-1}, \tag{9.2.27}$$

which is a few times the lattice constant even at $\alpha \sim 10$ which justifies the use of the continuum model (the upper limit k_0 of q -integration) up to this value of α .

The polaron energy in the strong coupling case such as (9.2.26) does not contain $\hbar \omega_{\ell}$ as is obvious from (9.1.15), being purely an electrostatic energy. This is because the angular velocity $\langle v_x \rangle / \langle r \rangle \sim (\alpha^2/4.5) \omega_{\ell}$ obtained from the orbital (9.2.25) is so much greater than ω_{ℓ} in the strong coupling regime that the lattice cannot follow the electronic motion, namely, the lattice is essentially static.

Although the calculation of the polaron mass in the strong coupling regime is possible, it is less reliable than that of the energy, and will be omitted here.

Feynman applied his path integral method (PIM) to the polaron problem,⁵ the method which had been presented by himself¹² as a spacetime approach to describe quantum mechanics. It is a kind of variational method which optimizes a trial model action for the electron moving in the phonon fields. It turned out to be more powerful than any other method available at that time over the entire range of the coupling constant.

Mishchenko *et al.*¹³ applied the recently-developed quantum Monte Carlo (QMC) method to the study of polarons, and showed that the calculation can be extended to any desired higher-order diagrams of polaron propagation dependent only on the cost of time on modern high-speed computers. Since the description of these two powerful methods in their technical details is beyond the scope of the

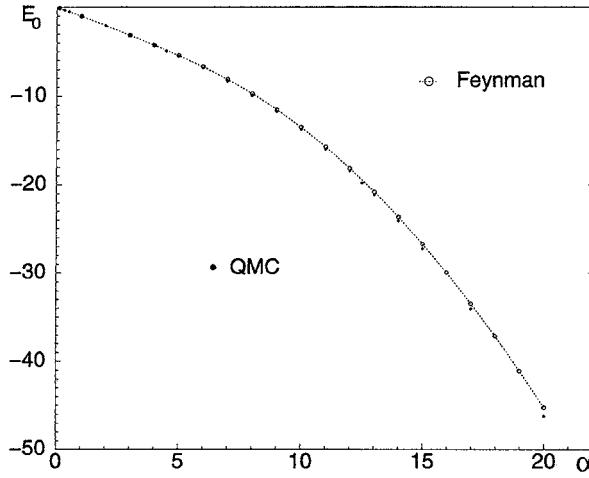


Fig. 9.1 Bottom of the polaron band as a function of α . Due to Mishchenko *et al.*¹³

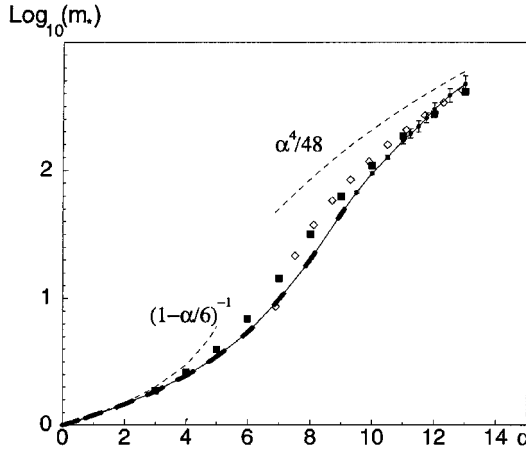


Fig. 9.2 Effective mass as a function of coupling constant α . Calculated values (solid line) are compared with perturbation theory and strong-coupling-limit results (dashed lines), Feynman's approach (filled squares), and the variational approach due to Feranchuk *et al.* (open diamonds). Due to Mishchenko *et al.*¹³

present book, we will present here only the results obtained by the latter study in comparison with those by the path integral and other methods.

The polaron energies at $K = 0$ as calculated by QMC (Mishchenko) and PIM (Feynman) are shown against the coupling constant α in Fig. 9.1, by the filled circles (the error is within the point) and open circles, respectively. One may be impressed how the variational method (PIM) gives values close to the exact method (QMC) over the entire range of α . The situation is somewhat different with the effective mass m_p as shown in Fig. 9.2, where not only the perturbation

theory and the strong coupling theory (both shown by broken lines) have large deviations in the intermediate coupling range, as they should, but also PIM and a variational calculation by Feranchuk *et al.*¹⁴ show significant deviations from QMC (solid lines, the error is smaller than the circle). The number of virtual phonons $\langle n \rangle$ calculated by QMC is shown by filled circles (relative accuracy better than 10^{-3}) in Fig. 9.3, in comparison with the intermediate coupling theory (dashed lines) and the parabolic interpolation of the strong coupling theory (solid line). An interesting feature appears with the spectral density $g_K(E)$, the weight of the bare electron with $K = 0$ contained in the polaron states with energy E , as shown in Fig. 9.4. In addition to the peak just beyond $E = 1 (\equiv \hbar\omega_\ell)$ which is expected from the perturbation theory), there appear extra peaks at ~ 3.5 and at ~ 8 whose heights

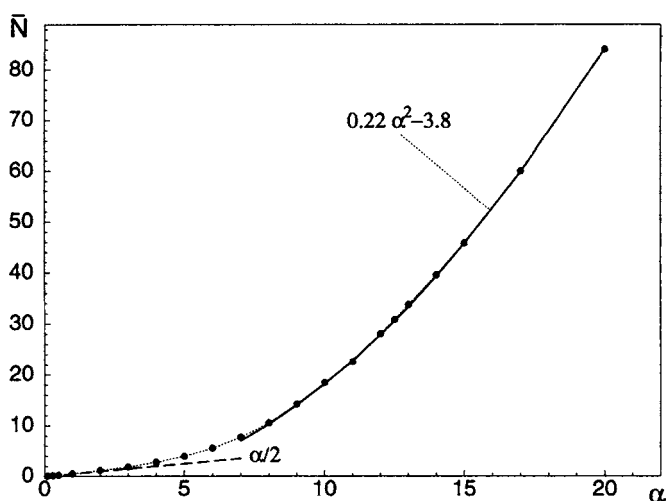


Fig. 9.3 The calculated average number of phonons (filled circles) in the polaron ground state as a function of α . The dashed line is the perturbation theory result, and the solid line is the parabolic fit for the strong coupling limit. Due to Mishchenko *et al.*¹³

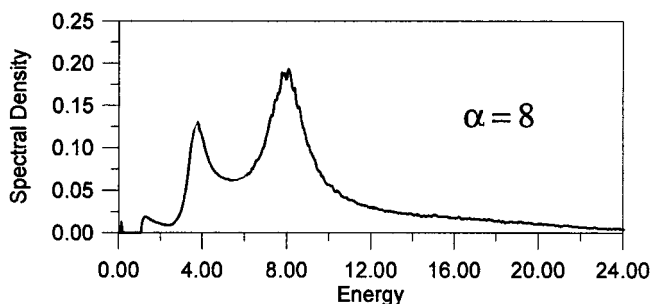


Fig. 9.4 Evolution of spectral density with α in the crossover region from intermediate to strong coupling. The energy is counted from the position of the polaron, in units of $\hbar\omega_\ell$. Due to Mishchenko *et al.*¹³

grow with increasing α without significant shifts in position. The physical origin of these peaks which are tentatively called “many-phonon unstable states” by the authors is not clear at the moment and is an interesting problem left for future study.

9.3 Electronic polaron, a renormalized electron in a many-electron field

We have so far considered the effect of the *low*-frequency components of electric polarization of an insulating crystal, the optical modes of lattice vibrations which contribute to the difference $\epsilon_e^{-1} - \epsilon_s^{-1}$, and studied how and to what extent they can follow a *slow* motion of an electron within a band. In this treatment, the high-frequency component of electric polarization due to bound electrons of the host crystal which contribute to the difference $\epsilon_0^{-1} - \epsilon_e^{-1}$ was considered to immediately follow the slow motion of the electron and was taken into account phenomenologically through the background dielectric constant ϵ_e . For the completion of the polaron theory from the dielectric viewpoint, it is now necessary to study how and to what extent the higher-frequency components of electric polarization can follow a faster motion of an electron within a band. The electron dressed with electronic polarization is called the “electronic polaron” in contradistinction to the ordinary polaron studied thus far. The latter may be called the “phonon polaron” when it is necessary to avoid confusion. In the phonon-polaron theory, it was implicitly assumed that “bare electron” means the electronic polaron as is obvious from the use of $\epsilon_e^{-1} - \epsilon_s^{-1}$ in the coupling constant.

The concept of the electronic polaron was first introduced¹⁵ with the use of a *single* longitudinal exciton band (in place of an optical phonon), the lowest branch of elementary excitation of the electronic system in an insulator. While a single longitudinal mode of the optical phonon is realistic, as is the case for isotropic diatomic crystals, a single longitudinal exciton band (used in early papers, for simplicity) is not, since there are a host of longitudinal exciton bands $E_{\lambda\ell K}$ of the Rydberg series (discrete quantum number, $\lambda = n$) followed by the ionization continuum (continuous quantum number, $\lambda = k$). Moreover, the transitions from all inner core bands below the valence band also make contributions (though small) to the electronic polarization and hence to $\epsilon_0^{-1} - \epsilon_e^{-1}$. Here all of these longitudinal electronic excitation modes (λ, \mathbf{q}) will be included in the Hamiltonian, which can therefore be written, in view of (9.1.4) and (9.1.7), as

$$\begin{aligned} H &= h_c + H_e + H_I, \quad h_c = p^2/2m_e, \quad H_I = \sum_{\lambda\mathbf{q}} E_{\lambda\mathbf{q}} b_{\lambda\mathbf{q}}^\dagger b_{\lambda\mathbf{q}}, \\ H_I &= (-e)\phi(\mathbf{r}) = V^{-1/2} \sum_{\lambda\mathbf{q}} [\partial\epsilon(\mathbf{q}, E_{\lambda\mathbf{q}})/\partial E_{\lambda\mathbf{q}}]^{-1/2} (ie/q) \\ &\quad \times [b_{\lambda\mathbf{q}} \exp(i\mathbf{q} \cdot \mathbf{r}) - b_{\lambda\mathbf{q}}^\dagger \exp(-i\mathbf{q} \cdot \mathbf{r})]. \end{aligned} \quad (9.3.1)$$

Here we have omitted the suffix ℓ and chosen $E_{\lambda\mathbf{q}} = \hbar\omega_{\lambda\mathbf{q}}$ as a variable in the dielectric function $\epsilon(\mathbf{q}, E_{\lambda\mathbf{q}})$. This is the realistic generalized version of the early

model with a single exciton band. The use of boson operators for excitons or electron-hole pairs is only for the sake of a compact expression and need not be justified since the coupling will turn out to be weak enough to allow perturbation theory and the expected number of electron-hole pairs to be smaller than unity (although the self-energy cannot be neglected).

Before studying the electronic polaron with the Hamiltonian (9.3.1), let us prove here a sum rule for the coupling coefficients in H_I which makes our problem transparent from the viewpoint of the dielectric theory. Confining ourselves to the long-wavelength limit ($\mathbf{q} \rightarrow 0$), we rewrite the reciprocal of (8.4.20), with ϵ_{ce} replaced by ϵ_0 (so as to include the core electrons), as

$$\epsilon_0/\epsilon(\omega) = \prod_{\lambda} (\omega_{\lambda t}^2 - \omega^2)/(\omega_{\lambda \ell}^2 - \omega^2) = 1 - \sum_{\lambda} F_{\lambda} \omega_{\lambda \ell}^2/(\omega_{\lambda \ell}^2 - \omega^2). \quad (9.3.2)$$

Differentiating with respect to ω^2 and comparing both sides at $\omega \sim \omega_{\lambda \ell}$ where the λ th term becomes dominant, one obtains the derivative:

$$[d\epsilon/d(\omega^2)]_{\lambda \ell} = \epsilon_0 [F_{\lambda} \omega_{\lambda \ell}^2]^{-1}. \quad (9.3.3)$$

Inserting this in (9.3.2) and putting $\omega = 0$ where $\epsilon(0) = \epsilon_e$ (see below), one obtains the sum rule:

$$\epsilon_0^{-1} - \epsilon_e^{-1} = \sum_{\lambda} [\omega^2 d\epsilon/d(\omega^2)]_{\lambda \ell}^{-1} = 2 \sum_{\lambda} [E d\epsilon/dE]_{\lambda \ell}^{-1}. \quad (9.3.4)$$

The contributions to the dielectric constant $\epsilon(\omega)$ from the bound electrons under consideration and from lattice vibrations studied in Section 5.3 are shown schematically in Fig. 5.3. As is obvious, the contributions from electrons deep in atomic core states must be considered in the summation on λ in (9.3.4) although they are smaller than those from valence-band electrons as will be mentioned in Chapter 14.

Coming back to (9.3.1), one can calculate the self-energy Δ_K of the electronic polaron by the second-order perturbation theory in the same way as (9.1.20) for the phonon polaron. Assuming parabolic forms for the conduction and exciton bands

$$\epsilon_c(\mathbf{k}) = \hbar^2 \mathbf{k}^2/2m_e, \quad E_{\lambda \ell} \mathbf{K} = E_{\lambda \ell} + \hbar^2 \mathbf{K}^2/2m_{\text{ex}}, \quad m_{\text{ex}} \equiv m_c + m_v, \quad (9.3.5)$$

and neglecting the \mathbf{q} -dependence of $\partial\epsilon(\mathbf{q}, E_{\lambda q})/\partial E_{\lambda q}$, one obtains

$$\begin{aligned} \Delta_K &= - \sum_{\lambda q} (\partial\epsilon/\partial E_{\lambda})^{-1} (e/q)^2 [E_{\lambda} + \hbar^2 \mathbf{q}^2/2m_{\text{ex}} + \hbar^2 \{(\mathbf{K} - \mathbf{q})^2 - \mathbf{K}^2\}/2m_e]^{-1} \\ &= -(e^2 k_0/2\pi^2) \sum_{\lambda} (E_{\lambda} \partial\epsilon/\partial E_{\lambda})^{-1} [P_{\lambda} + Q_{\lambda} \hbar^2 m_e' K^2/12m_e^2 E_{\lambda}], \end{aligned} \quad (9.3.6)$$

where $m_e'^{-1} \equiv m_e^{-1} + m_{\text{ex}}^{-1}$, $k_0 = (6\pi^2/v_0)^{1/3}$ (Debye cut-off for the q -integration) and

$$P_{\lambda} \equiv t_{\lambda}^{-1} \tan^{-1} t_{\lambda}, \quad Q_{\lambda} \equiv t_{\lambda}^{-1} \tan^{-1} t_{\lambda} - (1 - t_{\lambda}^2)(1 + t_{\lambda}^2)^{-2}, \quad (9.3.7)$$

$$\text{with } t_{\lambda} \equiv k_0/k_{\lambda}, \quad k_{\lambda} \equiv (2m_e' E_{\lambda}/\hbar^2)^{1/2}. \quad (9.3.8)$$

Denoting the average of P_λ (and other quantities with suffix λ) over λ , with relative weight $(E_\lambda \partial \epsilon / \partial E_\lambda)^{-1}$, by $\langle\langle P \rangle\rangle$, etc., one can rewrite (9.3.6) as

$$\Delta_K = -(\epsilon_0^{-1} - \epsilon_e^{-1})(e^2 k_0 / 4\pi^2) [\langle\langle P \rangle\rangle + \langle\langle Q/E \rangle\rangle \hbar^2 m'_e K^2 / 12m_e^2]. \quad (9.3.9)$$

The expected number of virtual electron-hole pairs in the electronic polaron with $K = 0$ is calculated as

$$\begin{aligned} n &= V^{-1} \sum_{\lambda q} (\partial \epsilon / \partial E_\lambda)^{-1} (e/q)^2 [E_\lambda + \hbar^2 q^2 / 2m'_e]^{-2} \\ &= (e^2 / 4\pi^2) (\epsilon_0^{-1} - \epsilon_e^{-1}) \langle\langle R E^{-1/2} \rangle\rangle, \end{aligned} \quad (9.3.10)$$

$$R_\lambda \equiv [\tan^{-1} t_\lambda + t_\lambda (1 + t_\lambda^2)^{-1}] / 2. \quad (9.3.11)$$

The electrostatic potential $\phi(\mathbf{r}')$ (see (9.3.1)) due to the electronic polarization caused by an electron at \mathbf{r}_e is calculated as

$$\begin{aligned} \langle \Psi(\mathbf{r}_e) | \phi(\mathbf{r}') | \Psi(\mathbf{r}_e) \rangle_{\text{field}} &= -V^{-2} \sum_{\lambda q} (\partial \epsilon / \partial E_\lambda)^{-1} (-e/q^2) \\ &\times [E_\lambda + \hbar^2 q^2 / 2m'_e]^{-1} [\exp\{i\mathbf{q} \cdot (\mathbf{r}' - \mathbf{r}_e)\} + \exp\{-i\mathbf{q} \cdot (\mathbf{r}' - \mathbf{r}_e)\}], \end{aligned} \quad (9.3.12)$$

where $\Psi(\mathbf{r}_e)$ represents the electronic-polaron wave function similar to (9.1.26) with $K = 0$ and $\langle \cdots \rangle_b$ indicates that the expectation value is taken as regards the field operators b and b^\dagger . Integrating the above expression on \mathbf{r}_e over the total volume V keeping $\mathbf{r}' - \mathbf{r}_e \equiv \mathbf{r}$ constant, one obtains the self-induced electrostatic potential $\phi_{\text{ind.}}(\mathbf{r})$ at distance \mathbf{r} from the electron:

$$\begin{aligned} \phi_{\text{ind.}}(\mathbf{r}) &= \int_v d\mathbf{r}_e \langle \Psi(\mathbf{r}_e) \Phi(\mathbf{r}_e + \mathbf{r}) | \Psi(\mathbf{r}_e) \rangle b \\ &= (e/\pi^2) \sum_{\lambda} (E_\lambda \partial \epsilon / \partial E_\lambda)^{-1} \int_0^{k_0} dk [\sin(kr)/(kr)] / [1 + (k/k_\lambda)^2]. \end{aligned} \quad (9.3.13)$$

Further study becomes somewhat complicated because of the λ -dependent quantities. Since the weight $(E_\lambda \partial \epsilon / \partial E_\lambda)^{-1}$ in taking their average seems to be more or less concentrated around the energy gap ε_g , we study the two limiting situations as defined below, and apply the results to a qualitative study of various materials which are somewhere between them. The first is the *classical* limit (c) with (most of) $t_\lambda = k_0/k_\lambda \ll 1$, namely, with the recoil kinetic energies of an electron and an exciton, even at their maximum values, negligible against the exciton energies E_λ and energy gap ε_g . It applies to wide-gap insulators. The second is the *quantal* limit (q) with (most of) $t_\lambda = k_0/k_\lambda \gg 1$ with the recoil kinetic energies being much greater than ε_g , which is applicable to narrow-gap semiconductors. Note that the phonon polaron studied in Sections 9.1 and 9.2 belongs to the (q) limit.

Let us first note that the integral in (9.3.13) can be replaced by

$$\int_0^\infty dk [\sin(kr)/(kr)] = (\pi/2)r^{-1}$$

provided that $k_0 r \gg 1$ in the (c) limit and $k_\lambda r \gg 1$ in the (q) limit since in either case (k/k_λ) can be neglected in the region of k important for the integral. With the use of (9.3.4), one obtains

$$\begin{aligned}\phi_{\text{ind.}}(\mathbf{r}) &= -(-e/2\pi) \sum_{\lambda} (E_{\lambda} \partial \epsilon / \partial E_{\lambda})^{-1} (1/r) \\ &= -(-e/4\pi) (\epsilon_0^{-1} - \epsilon_e^{-1}) (1/r).\end{aligned}\quad (9.3.14)$$

Adding this induced potential to the direct potential $\phi_{\text{dir.}}(\mathbf{r}) = -e/4\pi\epsilon_0 r$ of the electron, one finally obtains the screened potential:

$$\phi_{\text{scr.}}(\mathbf{r}) = \phi_{\text{ind.}}(\mathbf{r}) + \phi_{\text{dir.}}(\mathbf{r}) = (-e)/4\pi\epsilon_e r \quad (9.3.15)$$

at long distance, as is expected. This is the microscopic foundation for the use of the screened potential with the macroscopic dielectric constant ϵ_e for an electron with kinetic energy greater than the optical phonon energy but smaller than the energy gap of the crystal. The minimum distance down to which (9.3.15) holds is the radius of the electronic polaron which can be seen from the above argument and will be given explicitly later on.

Let us then study other quantities, starting with the (q) limit with $t_\lambda \gg 1$ which allows better comparison with the phonon polaron since the latter is also in the (q) limit. One can put $P_\lambda \sim Q_\lambda \sim (\pi/2)t_\lambda^{-1}$ in (9.3.9) and obtain

$$\Delta_K^{(q)} = -\alpha_{\text{ep}} \epsilon_g \{(\epsilon_g'/\epsilon_g)^{1/2}\}_{\text{I}} - (\alpha_{\text{ep}}/6) \{(m_e'/m_e)(\epsilon_g/\epsilon_g'')^{1/2}\}_{\text{II}} \hbar^2 K^2 / 2m_e, \quad (9.3.16)$$

$$\alpha_{\text{ep}}^{(q)} \equiv (1/2) [(\epsilon_0^{-1} - \epsilon_e^{-1})/4\pi] e^2 k_{\text{ep}} / \epsilon_g, \quad (9.3.17)$$

$$k_{\text{ep}}^{(q)} \equiv (2m_e' \epsilon_g / \hbar^2)^{1/2}, \quad (9.3.18)$$

$$\epsilon_g'^{1/2} \equiv \langle \langle E^{1/2} \rangle \rangle, \quad \epsilon_g''^{-1/2} \equiv \langle \langle E^{-1/2} \rangle \rangle. \quad (9.3.19)$$

Namely, with definitions (9.3.17, 9.3.18) of the electronic polaron which are obtained from those of the phonon polaron (9.1.15, 9.1.16) by replacing $\hbar\omega_l$ by ϵ_g , k_p by $k_{\text{ep}}^{(q)}$ and m_e by m_e' (the reduced mass of the conduction electron and the exciton). Equation (9.3.16) compares directly with (9.1.22) but for minor correction factors $\{\cdots\}_{\text{I}}$ and $\{\cdots\}_{\text{II}}$ for Δ_0 and effective mass change, respectively. The expected number of virtual electron-hole pairs are given, since $R_\lambda \sim \pi/4$, by

$$n^{(q)} = (1/2) \alpha_{\text{ep}}^{(q)} (\epsilon_g/\epsilon_g'')^{1/2}, \quad (9.3.20)$$

which is to be compared with (9.1.25) at $k = 0$. The induced electrostatic potential at $r = 0$ turns out to be

$$\phi_{\text{ind.}}^{(q)}(0) = -(-e)[k_{\text{ep}}/4\pi](\epsilon_0^{-1} - \epsilon_e^{-1})(\epsilon_g'/\epsilon_g)^{1/2} = 2\Delta_0/(-e). \quad (9.3.21)$$

One can define the radius r_{ep} of the electronic polaron by equating the long-distance formula (9.3.14) at $r = r_{\text{ep}}$ to $\phi_{\text{ind.}}(0)$ of (9.3.21):

$$r_{\text{ep}}^{(q)} = k_{\text{ep}}^{(q)-1} (\varepsilon_{\text{g}}/\varepsilon_{\text{g}}')^{1/2}, \quad (9.3.22)$$

which is to be compared with $r_{\text{p}} = k_{\text{p}}^{-1}$ for the phonon polaron (see (9.1.16) and the statement below (9.1.27)).

It should be noted here that for a quantitative study of narrow-gap semiconductors the non-parabolicity of the conduction and valence bands and the correction factors $(\varepsilon_{\text{g}}'/\varepsilon_{\text{g}}) > 1$ will play important roles.

In the opposite (c) limit, one notes $\langle\langle P \rangle\rangle = 1$ and $\langle\langle Q/E \rangle\rangle = 0$ in (9.3.7), whence one obtains

$$\Delta_{\text{K}}^{(c)} = -(1/2)(\epsilon_0^{-1} - \epsilon_{\text{e}}^{-1})(e^2/4\pi)(2k_0/\pi), \quad (9.3.23)$$

with no change in the effective mass m_{c} of the electron. The expected number of virtual electron-hole pairs is given by

$$n^{(c)} = |\Delta^{(c)}| \langle\langle 1/E \rangle\rangle \sim (|\Delta|/\varepsilon_{\text{g}})(\varepsilon_{\text{g}}/\varepsilon_{\text{g}}'). \quad (9.3.24)$$

The induced electrostatic potential at $r = 0$ is given by

$$\phi^{(c)}(0) = -(\epsilon_0^{-1} - \epsilon_{\text{e}}^{-1})(-e)k_0/2\pi^2. \quad (9.3.25)$$

By comparing this with the long-distance formula (9.3.14), one can define the radius of the electronic polaron in the (c) limit by

$$r_{\text{ep}}^{(c)} = \pi/2k_0. \quad (9.3.26)$$

Equations (9.3.22) and (9.3.26) give the lower bound of the validity range of the screened potential (9.3.15) used on a phenomenological argument. As is obvious from eqs. (9.3.23, 9.3.25 and 9.3.26) which contain neither \hbar nor ε_{g} , the electronic polaron in the (c) limit is almost a classical point charge imbedded within a medium of dielectric continuum except for the finite Debye cut-off k_0 or the lattice constant $a \sim \pi/k_0$.

It is also interesting to note that the bottom of the self-induced potential $(-e)\phi^{(c)}(0)$ is again twice $\Delta^{(c)}$, exactly the same factor as in the (q) limit. This is related to the fact that the self-energy can be written as

$$\Delta = -(1/2)(\epsilon_0^{-1} - \epsilon_{\text{e}}^{-1})(e^2/4\pi r_{\text{ep}}) \quad (9.3.27)$$

in both the limits, (q) and (c).

The fact that the coupling constant of the electronic polaron, α_{ep} in the (q) limit and $\Delta/\varepsilon_{\text{g}}$ in the (c) limit, and the expected number n of the virtual electron-hole pairs are a few tenths justifies the perturbation theory a posteriori. However, the sum of

the electron and hole self-energies (the latter may be larger in magnitude because usually $m_v > m_c$) due to the electronic-polaron effect amounts to a significant fraction of the energy gap calculated without consideration of such a correlation effect, as was studied more quantitatively and extensively for a variety of insulators by Fowler,¹⁶ Inoue, Mahutte and Wang,¹⁷ and Kunz and collaborators.¹⁸ Referring to a more recent study, the first-principle calculation of electronic states in one-dimensional polydiacetylene due to Suhai¹⁹ gave a bandgap of about 5.8 eV in the Hartree–Fock approximation, which was reduced to 3.7 eV by taking account of the correlation effect in terms of the electronic-polaron theory.

The screening of the Coulomb potential of a point charge $e/4\pi\epsilon_0 r$ in a vacuum into $e/4\pi\epsilon_e r$ in a dielectric medium with *electronic polarization only*, as described by (9.3.15), applies also to the effective potential between an electron and a hole in a dielectric medium, as was shown by Haken²⁰ by considering the electronic-polaron effect of both particles, with the validity range for r chosen to be greater than the sum of the radii of *electronic polarons* of the electron and the hole. He had also derived the effective potential in a dielectric medium with *both electronic and displacement polarizations* to be given by $e/4\pi\epsilon_s r$ with the validity range for r chosen to be greater than the sum of the radii of *phonon polarons* (described in Sections 9.1, 9.2) of the electron and the hole. In this case, the origin of the energy of the two particles at $r = \infty$ is at the bandgap which is renormalized *both by the electronic-polaron effects and the phonon-polaron effects* of the two particles. At long distance the screening factor is given by $1/\epsilon_s$, while at short distance it is given by $1/\epsilon_e$ since the displacement polarization (if any) can no longer follow the rapid relative motion of the polarons in this region, but the screened potential of the latter type must be measured from the bandgap renormalized by the *electronic-polaron effects only*.

A *virtual* exciton around the electron and the hole of a *real* exciton, as expected in the electronic-polaron theory, reveals itself as an exciton sideband of the real exciton, at around twice the exciton energy in the optical and the energy-loss spectra in alkali halides, as was pointed out by Miyakawa,²¹ Hermanson,²² and Devreese, Kunz and Collins.²³ Recently, the exciton sidebands of valence-electron excitation²⁴ and inner core excitation²⁵ in solid rare gases were unambiguously resolved from the background absorption by short-time excitation spectra of exciton luminescence.

From a more rigorous point of view, the electronic polaron must be considered self-consistently in the context of the many-body problem in the following sense. One starts from the field of electronic polarization based on given exciton and band structures and studies its effect on individual electrons. However, the renormalized energy of the electron must be incorporated into the exciton and band structures assumed on the starting point and the calculation of the individual energy must be repeated, and so on until everything converges to the same point, just

like the Hartree or Hartree–Fock calculations. Namely, the renormalized energies of individual electrons on the one hand and the exciton and band structures on the other hand must be solved in a self-consistent way. This amounts to an entire story of the many-body problem in a solid which has never been solved but which may turn out to play an important role in the electronic properties of known as well as unknown materials. The role of the electronic-polaron theory is to depict one aspect of the complicated many-body problem.

9.4 Short-range electron–phonon interaction and self-trapping

Throughout the study of the phonon polaron described in Sections 9.1 and 9.2, it has been repeatedly suspected that the electron–optical phonon interaction might have a certain threshold coupling constant beyond which the self-trapped state as predicted by Landau becomes more stable than the polaron state studied by Fröhlich and others. All studies of the effective mass of the phonon-dressed electron indicated a continuous, instead of an abrupt, transition from the almost free or polaron-like state to the almost self-trapped state. However, the self-trapped hole found in alkali halides by the electron spin resonance study due to Känzig *et al.*²⁶ showed a completely localized molecule-like structure, namely a positive hole shared by a pair of halide ions (X_2^-) which are closer than their normal distance as nearest neighbors. This self-trapped hole seemed almost immobile at low temperatures. It differed from what can be imagined for an electron in a dielectric continuum in its strong but realizable coupling constant, with respect to the effective mass and the microscopic structure.

The reason why the optical phonon was considered as the primary candidate responsible for self-trapping is that the direct Coulomb interaction between an electron (or a hole) and ions seemed to be of predominant importance in ionic crystals. Once incorporated into the conventional formalism of the electron–phonon interaction, this long-range interaction turned out to be more effective in its long-wavelength region, which justified the use of a dielectric continuum model *a posteriori*. It seems that this vicious circulation ruled out a possible role from short-range interaction.

However, it is not only the coupling strength but also the force range which characterizes the electron–phonon interaction. The acoustic mode of lattice vibrations, which is universal in all types of crystals – metallic, covalent, molecular as well as ionic – has a short-range interaction with an electron and, as will be shown below, this interaction behaves completely differently from the long-range interaction with the optical mode.

Consider an elastic continuum described by the field of an elastic strain tensor $e_{ij}(\mathbf{r})$, in place of the dielectric continuum described by the polarization field $\mathbf{P}(\mathbf{r})$.

The energy of the bottom of the conduction band will be subject to change which is a linear function of the components $e_{ij}(\mathbf{r})$ if they are small enough. In *isotropic* crystals, the conduction-band bottom, if situated at $\mathbf{k} = 0$, is subject to a local energy change at position \mathbf{r} which is given by

$$\Xi \sum_i e_{ii}(\mathbf{r}) \equiv \Xi [\partial \xi_x / \partial x + \partial \xi_y / \partial y + \partial \xi_z / \partial z] \equiv \Xi \Delta(\mathbf{r}) \quad (9.4.1)$$

in terms of a *single* coefficient Ξ and a *single* field variable $\Delta(\mathbf{r})$ due to the assumed symmetry of the crystal. Here, $\xi(\mathbf{r})$ and $\Delta(\mathbf{r}) \equiv \text{div } \xi(\mathbf{r})$ represent local displacement and local dilation, respectively, of the elastic continuum. An electron at or near the conduction-band bottom is subject to the potential given by (9.4.1) which is therefore called the “deformation potential”, and Ξ is called the deformation-potential constant, or simply deformation potential. It was first introduced by Bardeen and Shockley²⁷ to calculate the mobility (average velocity under unit electric field applied) of electrons in non-polar semiconductors. The value of Ξ is obtained by measuring the electron mobility, or by a quantum-mechanical calculation of the band energy as a function of the lattice constant. It is usually a few to several electron volts, being of the same order of magnitude as the bandwidth as is seen from its definition (9.4.1). (Note that Δ is a dimensionless quantity.) This is a typical *short-range* interaction of *contact type*, and is in contrast to the long-range point-dipole interaction between an electron and ionic displacements.

One can expand the displacement $\xi_n \equiv \xi(\mathbf{r} = \mathbf{R}_n)$ of the n th unit cell at \mathbf{R}_n in terms of the acoustic modes of lattice vibrations as

$$\xi_n = \sum_{sq} (\hbar/2NM\omega_{sq})^{1/2} \mathbf{u}_{sq} [b_{sq} \exp(i\mathbf{q} \cdot \mathbf{R}_n) + b_{sq}^\dagger \exp(-i\mathbf{q} \cdot \mathbf{R}_n)], \quad (9.4.2)$$

where M is the mass of a unit cell and \mathbf{u}_{sq} is a unit vector in the direction of displacement of the normal mode (sq), with $s = 1, 2, 3$ representing the three branches of the acoustic mode. The expression (9.4.2) can be derived from (5.5.2) with normalization condition (5.2.7) in the case of a monatomic lattice, and its generalization to a lattice consisting of polyatomic unit cells with the help of the relation $M_v^{-1/2} \zeta_v^{(sq)} \equiv \xi^{(sq)}$ (independent of v) $\equiv M^{-1/2} \mathbf{u}^{(sq)}$. In calculating $\text{div} \xi(\mathbf{r})$ with the use of (9.4.2), one again considers $\xi_n \equiv \xi(\mathbf{R}_n)$ as a continuous function of a continuous variable $\mathbf{r} = \mathbf{R}_n$. Since $\text{div}[\mathbf{u}_{sq} \exp(i\mathbf{q} \cdot \mathbf{r})] = i\mathbf{q} \cdot \mathbf{u}_{sq} \exp(i\mathbf{q} \cdot \mathbf{r})$, one finds that only the longitudinal component $\mathbf{q} \cdot \mathbf{u}^{(sq)}$ of each mode contributes to the deformation potential (9.4.1). Hence, with further approximation that the three acoustic modes consist of one purely longitudinal and two purely transverse modes (which holds only for particular directions of \mathbf{q} with high symmetry even in isotropic crystals), one can confine oneself to the longitudinal mode ($\mathbf{u}^{(sq)} \parallel \mathbf{q}$) dropping its suffix s and write down the potential energy (9.4.1) for an electron at the n th site

\mathbf{R}_n in the field of acoustic phonons as

$$\begin{aligned} \langle n | H_I | n \rangle &= \Xi \operatorname{div} \boldsymbol{\xi}(\mathbf{r} = \mathbf{R}_n) \\ &= \sum_{\mathbf{q}} \Xi (\hbar/2NM\omega_{\mathbf{q}})^{1/2} i\mathbf{q} [b_{\mathbf{q}} \exp(i\mathbf{q} \cdot \mathbf{R}_n) - b_{\mathbf{q}}^\dagger \exp(-i\mathbf{q} \cdot \mathbf{R}_n)]. \end{aligned} \quad (9.4.3)$$

We will adopt the Hamiltonian H_I with diagonal elements given by (9.4.3) and vanishing non-diagonal elements ($\langle n | H_I | n' \rangle = 0$ for $n \neq n'$) for the electron-acoustic phonon interaction.

In conformity with (9.4.3), the electron energy within the conduction band will also be written in the site representation as

$$\langle n | h_c | n' \rangle = t_{nn'} = t(\mathbf{R}_n - \mathbf{R}_{n'}) \quad (9.4.4)$$

(see Section 7.4). The corresponding \mathbf{k} -representation, the Fourier transform of (9.4.4), gives the dispersion of the conduction band:

$$\langle \mathbf{k} | h_c | \mathbf{k} \rangle = \varepsilon_c(\mathbf{k}) = \sum_{\mathbf{m}} t(\mathbf{R}_m) \exp(-i\mathbf{k} \cdot \mathbf{R}_m). \quad (9.4.5)$$

We will take the long-wavelength approximation for the angular frequency, $\omega_{\mathbf{q}} = c_\ell q$ with isotropic phase velocity c_ℓ of the longitudinal sound wave in the phonon Hamiltonian

$$H_L = \sum_{\mathbf{q}} \hbar \omega_{\mathbf{q}} b_{\mathbf{q}}^\dagger b_{\mathbf{q}} \quad (9.4.6)$$

and in the interaction Hamiltonian H_I . Since the interaction coefficient in (9.4.3) is proportional to $q^{1/2}$, in contrast to q^{-1} in the electron-optical phonon interaction (see (9.1.6)), the short-wavelength phonons play a dominant role here. This is the reason why we adopted site representation in (9.4.3) and (9.4.4) and considered the entire band structure (9.4.5) instead of expanding it in powers of $\mathbf{p} = \hbar\mathbf{k}$ (effective mass approximation (9.1.7)).

Another point of difference is the critical wave vector k_{ap} beyond which the electron can emit an acoustic phonon. Within the perturbation theory, it is given by $k_{\text{ap}} = 2m_e c_\ell / \hbar$ from the inequality $\hbar^2 k^2 / 2m_e > \hbar c_\ell k$. It is of the order of one-hundredth of the reciprocal lattice and one-tenth of k_p given by (9.1.16) since the sound velocity in a solid is $\sim 10^3 \text{ m s}^{-1}$. Namely, the region in which the electron is in a stationary state is confined to a very small part of the first Brillouin zone so long as the perturbation theory is valid.

Let us write the wave function for the stationary state of the electron-acoustic phonon system with total wave vector \mathbf{k} , the *acoustic polaron*, in a *most general* form *compatible with the translational symmetry* of the lattice,²⁸ as

$$\Psi_{\mathbf{k}} = (NS)^{-1/2} \sum_{\mathbf{m}} \exp(i\mathbf{k} \cdot \mathbf{R}_m) \psi_{\mathbf{k}}(\mathbf{m}) \chi_{\mathbf{k}}(\mathbf{m}) \quad (9.4.7)$$

where

$$\psi_k(\mathbf{m}) = \sum_n \phi_k(\mathbf{R}_n) |\mathbf{m} + \mathbf{n}\rangle \quad (9.4.8)$$

is the electron wave function with finite extension around the site \mathbf{m} , and $\chi_k(\mathbf{m})$ is the lattice wave function with finite extension of its distorted part around the site \mathbf{m} . Note that the vectorial notation \mathbf{m} is used to describe the three-dimensional lattice.

Let us apply the discrete lattice version of the Lee–Low–Pines variational method which amounts to choosing a trial wave function of the displaced lattice (without changing the frequencies ω_q) as shown by (9.2.6) to (9.2.8). In order to be compatible with the translational symmetry of the system, one puts the trial function as

$$\chi_k(\mathbf{m}) = \exp \left[\sum_q \left\{ f_k^*(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{R}_m) b_q - f_k(\mathbf{q}) \exp(-i\mathbf{q} \cdot \mathbf{R}_m) b_q^\dagger \right\} \right] |0\rangle \quad (9.4.9)$$

where $|0\rangle$ represents the phonon vacuum state. According to (9.2.8), the transformation $\exp[\cdot \cdot \cdot]$ in (9.4.9) has the effect of displacing the oscillator \mathbf{q} by

$$\Delta b_q = -f_k(\mathbf{q}) \exp(-i\mathbf{q} \cdot \mathbf{R}_m)$$

and hence the n th site atom by

$$\Delta \xi_n = - \sum_q (\hbar/2NM\omega_q)^{1/2} u_q [f_k(\mathbf{q}) \exp\{i\mathbf{q} \cdot (\mathbf{R}_n - \mathbf{R}_m)\} + (\text{c.c.})]$$

according to (9.4.2). The fact that it depends on \mathbf{n} only through $(\mathbf{R}_n - \mathbf{R}_m)$ assures the above-mentioned statement on $\chi_k(\mathbf{m})$ and hence the compatibility of Ψ_k given by (9.4.7) with translational symmetry of the system. The Ansatz (9.4.8) which is also necessary to assure compatibility replaces the role of the first transformation U_1 (see (9.2.2, 9.2.3) for consideration of electron–phonon recoil obeying the (pseudo-) momentum conservation rule which no more holds in the discrete lattice due to the Umklapp processes).

For the second step, we ignore possible recoil effects and put $\phi_k(\mathbf{R}_n) = \delta_{n0}$. Then Ψ_k of (9.4.7) satisfies the normalization condition: $\langle \Psi_k | \Psi_k \rangle = 1$ while the expectation value of the electronic energy (9.4.4), with only the nearest neighbor transfer $t(t_{nn'} = t$ when $\mathbf{R}_{n'} - \mathbf{R}_n = \boldsymbol{\delta}$, and $= 0$ otherwise, where $\boldsymbol{\delta}$ indicates the distance to any of the nearest neighbor sites) being considered for simplicity, turns out to be

$$\langle \Psi_k | h_c | \Psi_k \rangle = t \sum_{\boldsymbol{\delta}} \exp[i\mathbf{k} \cdot \boldsymbol{\delta} - \sigma_k(\boldsymbol{\delta})], \quad (9.4.10)$$

where $\boldsymbol{\delta}$ is to run over all nearest neighbors and

$$\sigma_k(\boldsymbol{\ell}) \equiv \sum_q |f_k(\mathbf{q})|^2 [1 - \exp(-i\mathbf{q} \cdot \boldsymbol{\ell})]. \quad (9.4.11)$$

The above results are obtained by noting the following expression for the overlap integral between the differently displaced lattice states:

$$\langle \chi_k(\mathbf{m}) | \chi_k(\mathbf{m} + \ell) \rangle = \exp[-\sigma_k(\ell)]. \quad (9.4.12)$$

Equation (9.4.12) is obtained by noting that the transformation $\exp[\sum_q \{\dots\}]$ of (9.4.9) giving χ_k consists of the product: $\Pi_q \{\dots\}$ due to the commutativity of the b_q s with different q s. The calculation of (9.4.12) then reduces to that of the following type of expression for each q :

$$F(\lambda) \equiv \langle 0 | \exp(-\lambda A) \exp(\lambda A') | 0 \rangle \quad (9.4.13)$$

with

$$A = \alpha^* b - \alpha b^\dagger, \quad A' = \alpha'^* b - \alpha' b^\dagger. \quad (9.4.14)$$

One can verify that the derivative of $F(\lambda)$ is

$$\begin{aligned} F'(\lambda) &= \langle 0 | \exp(-\lambda A) (-A) \exp(\lambda A') | 0 \rangle + \langle 0 | \exp(-\lambda A) A' \exp(\lambda A') | 0 \rangle \\ &= \{\alpha^*(\alpha' - \alpha) - \alpha'(\alpha'^* - \alpha^*)\} \lambda F(\lambda) \end{aligned}$$

by commuting the b operator to the right and the b^\dagger operator to the left. Integrating one obtains

$$F(\lambda) = \exp[\{\alpha^*(\alpha' - \alpha) - \alpha'(\alpha'^* - \alpha^*)\} \lambda^2 / 2], \quad (9.4.15)$$

whence follows (9.4.12) with (9.4.11).

The expectation values of the interaction and phonon Hamiltonians are readily calculated as

$$\langle \Psi_k | H_I | \Psi_k \rangle = - \sum_q E_d (\hbar / 2NM c_\ell)^{1/2} i q^{1/2} \{f_k(\mathbf{q}) - f_k^*(\mathbf{q})\}, \quad (9.4.16)$$

$$\langle \Psi_k | H_L | \Psi_k \rangle = \sum_q \hbar c_\ell q |f_k(\mathbf{q})|^2. \quad (9.4.17)$$

Adding (9.4.10), (9.4.16) and (9.4.17) and minimizing with respect to the trial function $f_k(\mathbf{q})$ gives

$$\begin{aligned} f_k(\mathbf{q}) &= -i q^{1/2} \Xi (\hbar / 2NM c_\ell)^{1/2} \\ &\times \left[-t \sum_\delta \{1 - \exp(i\mathbf{q} \cdot \delta)\} \exp\{i\mathbf{k} \cdot \delta - \sigma_k(\delta) + \hbar c_\ell q\} \right]_I^{-1}. \end{aligned} \quad (9.4.18)$$

Substituting this into (9.4.11) gives a set of ν simultaneous equations for the σ_s (where ν is the number of nearest neighbors)

$$\sigma_k(\delta) = \sum_q \Xi^2 (\hbar / 2NM c_\ell) q [1 - \exp(i\mathbf{q} \cdot \delta)] [\dots]_I^{-2}, \quad (9.4.19)$$

where $[\cdot \cdot \cdot]_I$ is the same energy denominator as appeared in (9.4.18). The corresponding value of the total energy is given by

$$E(\mathbf{k}) = t \sum_{\delta} [1 + \sigma_{\mathbf{k}}(\delta)] \exp[i\mathbf{k} \cdot \delta - \sigma_{\mathbf{k}}(\delta)] - \sum_{\mathbf{q}} \Xi^2 (\hbar/2NM_{C\ell})q [\cdot \cdot \cdot]_I^{-1}. \quad (9.4.20)$$

We study the behavior of this function in two limiting cases:

(i) Ξ small. In this case $f_{\mathbf{k}} \propto \Xi$ and $\sigma_{\mathbf{k}} \propto \Xi^2$. Noting that the band energy in the nearest neighbor transfer t (assumed hereafter to be negative for the band bottom to be at $\mathbf{k} = 0$) is given by

$$\varepsilon_c(\mathbf{k}) = -|t| \sum_{\delta} \exp(-i\mathbf{k} \cdot \delta), \quad (9.4.21)$$

one can write the total energy as

$$E(\mathbf{k}) = \varepsilon_c(\mathbf{k}) - \sum_{\mathbf{q}} \Xi^2 (\hbar/2NM_{C\ell})q / [\varepsilon_c(\mathbf{k} - \mathbf{q}) + \hbar c_{\ell}q - \varepsilon_c(\mathbf{k})] + O(\Xi^4). \quad (9.4.22)$$

The second term is the perturbation theory expression for the self-energy of the electron in the phonon field.

(ii) Ξ large. In this case $\sigma_{\mathbf{k}} \gg 1$ and the total energy reduces to

$$E(\mathbf{k}) = -\Xi^2/2M_{C\ell}^2 \equiv -E_R. \quad (9.4.23)$$

It represents the self-trapped state with infinite effective mass. E_R represents the energy of the *lattice relaxation* due to the localized electron, and corresponds to the energy E_R of *molecular relaxation* after optical excitation. Note that the band energy $\varepsilon_c(\mathbf{k})$ and the transfer energy t do not play any role in (9.4.23), suggesting that the two states (i) and (ii) are completely different in nature.

It is therefore of interest to study how the two states are connected as the coupling constant increases continuously. To see the essential features we approximate the band energy by a parabolic form

$$\varepsilon_c(\mathbf{k}) - \varepsilon_c(0) = -|t| \sum_{\delta} [\exp(-i\mathbf{k} \cdot \delta) - 1] \cong \hbar^2 k^2 / 2m_e, \quad (9.4.24)$$

and solve (9.4.19) for $\mathbf{k} = 0$. The equations will be expressed in terms of two dimensionless parameters:

$$\gamma \equiv \hbar c_{\ell} k_0 / (\hbar^2 k_0^2 / 2m_e), \quad (9.4.25)$$

$$g \equiv E_R / \nu |t|, \quad (9.4.26)$$

where k_0 is the Debye cut-off wave number defined below (9.3.6). We shall call g the coupling constant and γ the non-adiabaticity parameter. The parameter γ is the ratio of the maximum phonon energy to the bandwidth and is usually of the order of 10^{-2} to 10^{-1} . It compares the rapidity of phonon motion to that of electron motion, and

as $\gamma \rightarrow 0$ one obtains adiabatic conditions. In terms of these parameters, one can write eqs. (9.4.19, 9.4.20) as

$$\sigma_0(\delta) \equiv \sigma = 3g\gamma \exp(2\sigma) \int_0^1 dt t^3 [t + \gamma \exp(\sigma)]^{-2}, \quad (9.4.27)$$

$$-E(0)/\nu|t| = (1 + \sigma) \exp(-\sigma) + 3g\gamma \exp(\sigma) \int_0^1 dt t^2 [t + \gamma \exp(\sigma)]^{-1}. \quad (9.4.28)$$

Instead of eliminating σ , it is more convenient to take the parametric representation of E and g , with σ as a parameter. When $\gamma \exp(\sigma) \ll 1$, we have

$$g^{-1} \sim 3\gamma \exp(\sigma)/2\sigma, \quad \text{and} \quad -E(0)/\nu|t| \sim (1 + 2\sigma) \exp(-\sigma)$$

while, when $\gamma \exp(\sigma) \gg 1$,

$$g^{-1} \sim 3/4\gamma\sigma, \quad \text{and} \quad -E(0)/\nu|t| \sim 4\gamma\sigma/3.$$

The behavior of these two functions is shown in Fig. 9.5. The coupling constant is not a monotonic function of σ ; its actual functional dependence leads to the formation of a closed loop in the E - g curve as shown in Fig. 9.6. As the loop is always above the rest of the curve, the lower part of the loop represents a metastable state; the upper part (that is the solution with the highest energy) seems to correspond to the saddle point, or the activated state through which the stable state is connected

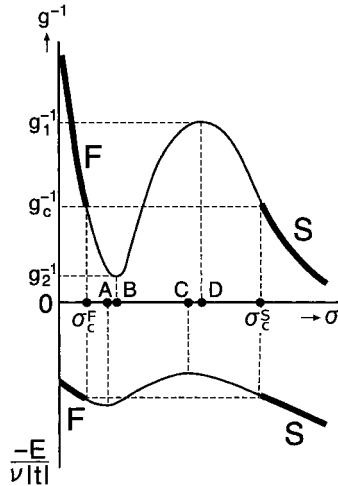


Fig. 9.5 Coupling constant g and calculated energy E , plotted against parameter σ , which is a measure of the distortion.²⁷

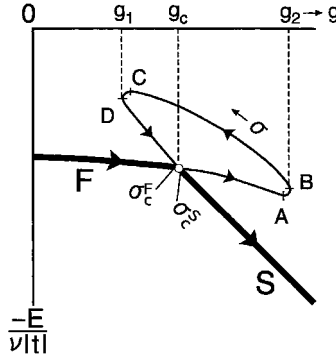


Fig. 9.6 Plot of E as a function of g . The arrows on the curves are in the direction of increasing σ .²⁷

with the metastable state. The stable part of the curve exhibits a discontinuous transition from the nearly free electron (moving with small distortion through the lattice) to the self-trapped state; i.e., σ changes discontinuously from $\sigma_c^F \sim 3\gamma/2 (\ll 1)$ to $\sigma_c^S \sim 3/4\gamma (\gg 1)$ at this kink point which corresponds to $g_c = 1 + 0(\gamma)$. Now σ , given by (9.4.11) with $\ell = \delta$, is related to the number of virtual phonons around the electron,

$$n = \langle \Psi_k | \sum_q b_q^\dagger b_q | \Psi_k \rangle = \sum_q |f_k(q)|^2.$$

The ground state includes very few ($\ll 1$) or a great number ($\gg 1$) of phonons according as $g < g_c$ or $g > g_c$. One can define the effective mass of the dressed electron from the k dependence of the total energy (9.4.20). The weak coupling solution has an effective mass m^* almost equal to m_c , while in the strong coupling solution $m_p/m_e = \exp(\sigma)$, which is an enormous quantity, since $\sigma > 3/4\gamma$ for $g > g_c$, i.e., σ is of the order of 10 to 100. It is then almost meaningless to speak of the self-trapped electron forming a band. Because of this large effective mass, the critical wave vector $k_{ap} (\sim 2m_p c_\ell / \hbar)^{1/2}$ beyond which the electron can emit an acoustic phonon is supposed to be too large to be meaningful in the strong coupling case. The situation is quite different for the case of the polaron, where neither the discontinuity nor such a large distortion is expected to take place.

It should be emphasized that the large discontinuity and the coexistence of essentially different states are due to the smallness of the non-adiabaticity parameter γ . In fact, if γ is larger than a certain critical value γ_0 (of the order of unity) the curve of g^{-1} against σ becomes monotonic and the discontinuity disappears. The situation is analogous to the problem of a gas-liquid transition, if we let γ correspond to temperature.

We must now discuss the approximation we have made, especially our Ansatz, $\phi(\mathbf{R}_n) = \delta_{n0}$. If we treat the polar mode in the same way, with the same Ansatz, there appears a discontinuity, somewhat less significant but of the same kind as was obtained here. However, this discontinuity is an artefact of the above-mentioned Ansatz. That the Ansatz $\phi(\mathbf{R}_n) = \delta_{n0}$ is inappropriate for the polar mode but appropriate for the acoustic mode will be shown from a different viewpoint in Section 9.5.

Although the above statements based on the variational method are supposed to be essentially correct from the physical viewpoint, one must be more careful about the nature of what we called the discontinuous change of the electron–phonon state against the continuous change of the coupling constant, as viewed with mathematical rigor. A real discontinuity can appear only in the adiabatic limit, $\gamma \rightarrow 0$. Let us consider the problem in a more practical way. Even if the two states, stable and metastable, coexist in a certain region of the coupling constant within the variational method with the specified form of trial wave function as we have mentioned, they are not exactly orthogonal to each other. If we were to consider a linear combination of the almost free and almost self-trapped states and minimize the total energy with respect to the coefficients of the linear combination, we would find a mathematically continuous but physically abrupt change of the nature (inclusive of the effective mass) of the true ground state against a continuous change of the coupling constant. A simple estimation of this abruptness is as follows. The tunneling energy between the two states mentioned above is of the order of $\nu|t| \exp(-\sigma)$ while their energy difference $\sim E_R - \nu|t|$ changes as $(g - g_c)\nu|t|$. Therefore, the crossover between the two states takes place practically within an extremely small range of $\Delta g \sim \exp(-\sigma)$. This enormous rapidity might be significantly reduced if one relaxes the constraint of $\phi(\mathbf{R}_n) = \delta_{n0}$. How to relax the constraint and to what extent the rapidity is reduced is not known. An important problem is to study how the rapidity of change with coupling constant depends on γ . This problem was studied by Sumi²⁹ with the use of the dynamical coherent potential method, a version of the coherent potential approximation for static fluctuation (Section 8.7) extended to dynamical fluctuation such as lattice vibrations. The dispersionless phonon model adopted because of technical reasons of calculation, which can be usefully applied to non-polar optical modes in molecular crystals, is supposed to underestimate the rapidity of change mentioned above, to which the long-wavelength acoustic phonons might contribute more because of their greater adiabaticity.

9.5 Continuum model for self-trapping

We learned in the preceding section that an electron in the phonon field may have two different states (one stable and the other metastable) in the adiabatic limit

($\gamma \rightarrow 0$) corresponding to different local minima of the adiabatic potential. The lifetime of the metastable state is proportional to the reciprocal of its quantum-mechanical overlap integral with the stable state ($\sim \exp(-\sigma')$, $\sigma' \sim \gamma^{-1}$), and may be quite long. It is instructive to study the possible extrema of the adiabatic potential of an electron in a deformable lattice, confining oneself to the adiabatic limit.

The entire lattice consists of $3N$ configuration coordinates where N is the total number of atoms contained in it, an intractably large number. However, our primary concerns are the qualitative aspects: which of the free state or the self-trapped state does the electron prefer, what is the spacial extension of the latter (namely, how to relax the restrictive Ansatz $\phi_k(\mathbf{R}_n) = \delta_{n0}$, how do the different force ranges of the optical and acoustic mode interactions feature in the adiabatic potential and what is the effect of coexistence of the optical and acoustic modes, and, finally, what is the role of the dimensionality of the lattice? To answer these questions, it is enough to treat the lattice as a continuous medium and the Bloch electron as a free electron with effective mass m_c as the only parameter. The optical and acoustic modes of the lattice vibrations can then be described in terms of a continuum with an appropriate dielectric constant ϵ (given by $\epsilon^{-1} = \epsilon_e^{-1} - \epsilon_s^{-1}$, as will be seen later) and an appropriate elastic constant C . Putting aside the kinetic energy (K_L) of the lattice, one can write the energy of the interacting system of an electron and the medium as a functional of the electron wave function $\psi(\mathbf{r})$, the dilation field $\Delta(\mathbf{r})$ and the electrostatic potential $\Phi(\mathbf{r})$ due to the displacement polarization $\mathbf{P}(\mathbf{r})$ of the medium,³⁰ as

$$\begin{aligned} W[\psi, \Delta, \Phi] = & (\hbar^2/2m_c) \int d\mathbf{r} (\nabla\psi)^2 \\ & + \Xi \int d\mathbf{r} \psi(\mathbf{r})^2 \Delta(\mathbf{r}) + (-e) \int d\mathbf{r} \psi(\mathbf{r})^2 \Phi(\mathbf{r}) \\ & + (C/2) \int d\mathbf{r} \Delta(\mathbf{r})^2 + (\epsilon/2) \int d\mathbf{r} [\nabla\Phi(\mathbf{r})]^2. \end{aligned} \quad (9.5.1)$$

The first term on the r.h.s. represents the electron kinetic energy (one-electron energy h_c measured from the band bottom), the second and third the interactions (H_I) of the electron with the deformation potential (see (9.4.1) and the electrostatic potential, respectively, and the fourth and fifth the potential energies (U_L) of the dilation and polarization fields, respectively. The longitudinal modes of the acoustic and optical vibrations are now replaced by the fields $\Delta(\mathbf{r})$ and $\nabla\Phi(\mathbf{r})$, respectively, of a continuous variable \mathbf{r} .

By minimizing (9.5.1) with respect to ψ , one would obtain the adiabatic potential $W[\Delta, \Phi]$ for the lowest electronic state in the ∞ -dimensional configuration coordinate space ($\Delta(\mathbf{r})$, $\Phi(\mathbf{r})$). This will be done in a simplified form later on. On the other hand, if one is interested only in the extrema of the adiabatic potential $W[\Delta, \Phi]$,

one can make an alternative but simpler approach by inverting the order of extremization, following Kubo³¹ who first considered the electron–polarization interaction by the field theoretical picture. By first minimizing (9.5.1) with respect to Δ and Φ , one obtains $\Delta(\mathbf{r}) = -(\Xi/C)\psi(\mathbf{r})^2$ and $\nabla^2\Phi(\mathbf{r}) = -(-e/\epsilon)\psi(\mathbf{r})^2$. (By comparing the solution of the latter equation at long distance, $\Phi(\mathbf{r}) = -(-e)/4\pi\epsilon r$ with the self-induced potential (9.1.27) of the polaron at long distance, one can confirm that $\epsilon^{-1} = \epsilon_e^{-1} - \epsilon_s^{-1}$ as mentioned above.) Putting them back into (9.5.1) gives

$$W[\psi] = (\hbar^2/2m_e) \int d\mathbf{r} (\nabla\psi)^2 - (1/2) \iint d\mathbf{r} d\mathbf{r}' \psi(\mathbf{r})^2 [(\Xi^2/C)\delta(\mathbf{r} - \mathbf{r}') + (e^2/4\pi\epsilon)|\mathbf{r} - \mathbf{r}'|^{-1}] \psi(\mathbf{r}')^2. \quad (9.5.2)$$

as a functional of ψ . The second line represents the self-interaction, namely, the electron-induced lattice distortions acting back upon the electron. It consists of -1 from H_I (the second and third terms of (9.5.1) which is linear in the fields) and $+(1/2)$ from U_L (the fourth and fifth terms which is quadratic), being an example of the virial theorem. The different forms of the kernels of the self-interaction originate from the different force range – the short-range electron dilation interaction and the long-range electron polarization interaction.

For the second step, we have to find the extrema of the functional (9.5.2) with respect to the wave function $\psi(\mathbf{r})$. Choosing the trial function $\psi(\mathbf{r}) = (2/a^2)^{3/4} \times \exp[-\pi(r/a)^2]$ with variational parameter a , one obtains

$$W[\psi] = W(a) = 3\pi\hbar^2/2m_e a^2 - \Xi^2/2C a^3 - e^2/4\pi\epsilon a. \quad (9.5.3)$$

Since we are based on the effective mass theory for the electron, $\varphi(\mathbf{r})$ here represents the envelope function $F(\mathbf{R}_n)$ of Section 7.4. Hence the orbital radius a in the electron wave function cannot be taken smaller than the lattice constant, a_0 . For this reason, it is more convenient to rewrite (9.5.3) as

$$W(a) \equiv E(\lambda) = B\lambda^2 - E_R^{(\text{ac})}\lambda^3 - E_R^{(\text{op})}\lambda \quad (9.5.4)$$

in terms of the degree of localization λ defined by

$$\lambda \equiv a_0/a \quad (0 < \lambda < 1), \quad (9.5.5)$$

where

$$B \equiv 3\pi\hbar^2/2m_e a_0^2, \quad E_R^{(\text{ac})} \equiv \Xi^2/2C a_0^3, \quad E_R^{(\text{op})} \equiv e^2/4\pi\epsilon a_0 \quad (9.5.6)$$

represent, respectively, the electron kinetic energy and the lattice relaxation energies due to the acoustic and optical modes, at complete localization: $\lambda = 1$.

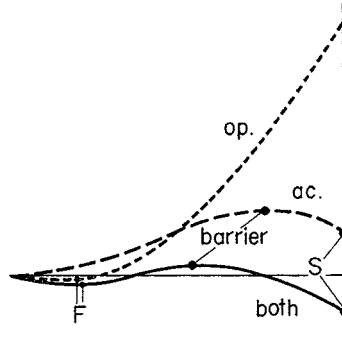


Fig. 9.7 Adiabatic energy of an electron in a deformable lattice as a function of the degree of localization λ .

As a function of λ ($0 < \lambda < 1$), the energy (9.5.4) has one or two minima as shown schematically in Fig. 9.7, depending on the values of the short- and long-range electron–phonon coupling constants defined by

$$g_s \equiv E_R^{(ac)}/B, \quad g_\ell \equiv E_R^{(op)}/B. \quad (9.5.7)$$

The first minimum F at $\lambda_F (< 1)$ represents a relatively delocalized state accompanied by moderate polarization but little contraction (or dilation depending on the sign of Ξ), while the second minimum S at $\lambda_S (=1)$ corresponds to a completely localized state (self-trapped state) with strong contraction (dilation) as well as strong polarization. The maximum between them represents the energy barrier (saddle point of the adiabatic potential) separating the states F and S. The height of the energy barrier for the $F \rightarrow S$ transition amounts to

$$\varepsilon_b = W_M - W_F = (4/27)Bg_s^{-2}(1 - 3g_sg_\ell)^{3/2}, \quad (9.5.8)$$

as is readily derived from (9.5.4, 9.5.7). Figure 9.8 is the *phase diagram* (different from the conventional usage of this terminology) for the *stable* and *metastable* (shown in parenthesis) states on the (g_s, g_ℓ) -plane. The thick solid line, across which the stable state changes abruptly from F to S, starts from the g_s -axis (at $g_s = 1$) and terminates at $g_\ell = 1, g_s = 1/3$ without reaching the g_ℓ -axis; obviously the only trigger for the F–S discontinuity is the short-range interaction, while the F-state is dominated by the long-range interaction with optical phonons.

It should be noted that the polaron state of the weak through strong coupling regime described in Sections 9.1 and 9.2 corresponds to the F-state of the above phase diagram along the g_ℓ -axis (note that $\alpha \sim g_\ell(B/\hbar\omega_\ell)^{1/2}$). The fact that our solution for the F-state apparently corresponds to the strong coupling polaron of the Pekar type is an artefact of our adiabatic description ($\omega_\ell \rightarrow 0$ so that $\alpha \rightarrow \infty$ in spite of finite g_ℓ). Not only the weak coupling but also the strong coupling polaron

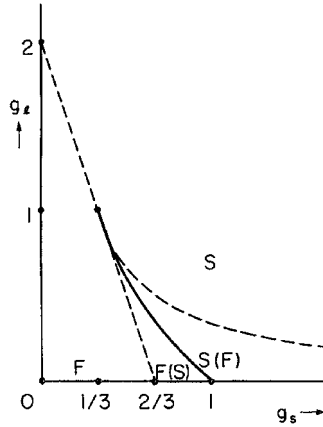


Fig. 9.8 Phase diagram for the stable and metastable (in parentheses) states where g_s and g_l denote the electron–phonon coupling constants of short and long ranges, respectively.³⁰

should be well distinguished from the S-state. The latter wants to be as small as possible ($a = a_0$) while the former has an optimum radius which is usually much greater than a_0 . For this reason, F- and S-states are sometimes called “large” and “small” polarons, respectively. This nomenclature, although convenient, should be used with some caution since it may give the impression that the latter is simply the strong coupling limit of the electron–polarization (optical phonon) interaction. The real S-state, which is well distinguished from, and can coexist with, the F-state, is triggered by the short-range interaction with acoustic phonons, although optical phonons make a comparable contribution to this state. The importance of the short-range interaction for self-trapping was first pointed out by Rashba.³² See also his review article on self-trapping of excitons.³³

For the $d(=1, 2 \text{ and } 3)$ -dimensional lattice, the second term (short-range interaction) in (9.5.3) turns out to be proportional to a^{-d} . In the absence of the long-range interaction whose dependence on dimensionality cannot be defined unambiguously, this short-range interaction in a one-dimensional lattice plays the same role as the long-range interaction did in a three-dimensional lattice, resulting in the absence of a discontinuity, as was pointed out by Emin and Holstein.³⁴ The situation in a two-dimensional lattice is peculiar and critical: the stable state is either completely free without any distortion ($\lambda = 0$) or completely localized state ($\lambda = 1$) according as $g_s < \text{or } > 1$, there being no metastable state under any situation. See, however, the adiabatic approach given later on in this section.

The discontinuity line in Fig. 9.8 indicates that the stable states of the electronic carriers – electrons, holes or excitons – in insulators are distinctly classified into F- and S-types, except in extremely narrow bands such that $g_l \sim \alpha(\hbar\omega_l/B)^{1/2} > 1$. On both sides of the discontinuity line, there are regions with coexistent stable and

metastable states separated by an adiabatic potential barrier. This is borne out by a variety of experimental studies on different types of materials as will be described later in this section.

While we have so far resorted to the inverted order of minimization of total energy, we now return to the normal order and study the adiabatic potential as a functional of the dilation field, ignoring the polarization field since our purpose is to study the multistability in the configuration coordinate space which arises only in the former field. We will extend our study to lower-dimensional systems, partly to elucidate the role of dimensionality for multistability and partly for application to such material systems which are increasingly coming to reality in these days.

Although it is intractable to deal with infinite-dimensional configuration coordinate (C.C.) space of dilation field $\Delta(\mathbf{r})$, it is enough, for our purpose of characterizing the C.C. path to self-trapping, to consider the two-dimensional C.C. space (Δ, a) , the magnitude and the spacial extension of the dilation $\Delta(\mathbf{r})$.²⁷ For simplicity, we put $\Delta(\mathbf{r}) = \Delta$ for $r < a$ and $\Delta(\mathbf{r}) = 0$ for $r > a$. Then the Schrödinger equation for the electron can be written, with suitable scaling, as

$$[-\Lambda^{-1}\nabla_{\rho}^2 - u(\rho)]\psi(\rho) = [E(\Delta, a)/(-\Xi\Delta)]\psi(\rho) \quad (9.5.9)$$

where $\rho \equiv r/a$, $u(\rho) = 1$ or 0 according as $\rho < \text{or} > 1$,

$$\Lambda \equiv -\Xi\Delta/K, \quad K \equiv \hbar^2/2m_e a^2 \quad (9.5.10)$$

and $E(\Delta, a)$ denotes the lowest eigenenergy of the electron, be it bound or free state. In terms of the depth, $-\Xi\Delta$, of the square well potential, one can write the binding energy as

$$-E(\Delta, a) = (-\Xi\Delta)f_d(\Lambda). \quad (9.5.11)$$

The function $f_d(\Lambda)$ calculated for dimensionalities $d = 1, 2, 3$ are shown in Fig. 9.9, the features of which are well known in the quantum mechanics of a square well potential³⁵ and are essentially the same in the discrete lattice.³⁶

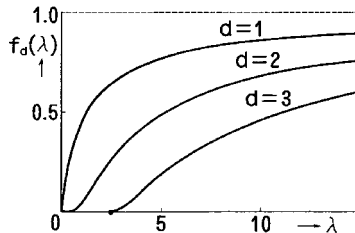


Fig. 9.9 The ratio binding energy to square well potential depth as a function of the latter, in 1-, 2-, and 3-dimensional lattices.³⁶

By adding U_L , one obtains the adiabatic energy

$$W(\Delta, a) = \gamma_d(C/2)a^d \Delta^2 - (-\Xi\Delta)f_d(\Lambda) \quad (9.5.12)$$

where γ_d is the volume of the d -dimensional sphere of unit radius. For one- and three-dimensional cases, one can bring (9.5.12) into the parameterless form

$$w(x, y) = \lambda^{d/2} x^{2-(d/2)} - f_d(\Lambda)x \quad (\Lambda = x/y) \quad (9.5.12a)$$

by rescaling the energy in units of W_0 :

$$-\Xi\Delta = W_0x, \quad K = W_0y, \quad W = W_0w, \quad (9.5.13)$$

$$(2m_e W_0/\hbar^2)^{(d/2)-1} \equiv \gamma_d \hbar^2 C/4m_e \Xi^2 \quad (d = 1, 3). \quad (9.5.14)$$

This procedure is impossible with $d = 2$, the marginal dimensionality; for any choice of W one obtains

$$w(x, y) = G^{-1}\lambda x - f_2(\Lambda)x \quad (d = 2) \quad (9.5.12b)$$

with an irremovable coupling constant

$$G \equiv 4m_e \Xi^2/\pi \hbar^2 C \quad (d = 2). \quad (9.5.15)$$

The adiabatic potentials for $d = 1, 2$ and 3 systems calculated with the use of (9.5.12a, 9.5.12b) and $f_d(\Lambda)$ (Fig. 9.9) are shown on the x (lattice distortion) – y (inverse square of the radius of distorted region) plane in Fig. 9.10 (a, b, b' and c), respectively,³⁶ where the negative energy region is shown by hatching. Noting that the volume $\gamma_d a^d$ of the distorted region should not be smaller than the atomic volume a_0^d of the discrete lattice, one finds the upper bound

$$y_d^{(u)} = \gamma_d^{4/(2d-d^2)} (8dg)^{-2/(2-d)} \quad (d = 1, 3) \quad (9.5.16)$$

for the physically meaningful y range where use has been made of (9.5.13, 9.5.14) and the relations $E_R = \Xi^2/2Ca_0^d$ and

$$\hbar^2/2m_e a_0^2 = |t| = B/2d. \quad (9.5.17)$$

With increasing g , the critical line $y = y_d^{(u)}$ moves downward for $d = 1$ and upward for $d = 3$.

Note in Fig. 9.10(a) that the free state ($\Delta = 0$) in a one-dimensional lattice is always unstable against lattice distortion Δ and there is only a minimum indicated by \times . This means that as g increases the optimum radius of the distorted region and the electron orbital decreases as g^{-1} while the optimum distortion $|\Delta_m|$ and the stabilization energy $-W_m$ increases as $g^{3/2}$ and g^2 , respectively (9.5.10, 9.5.13, 9.5.14), always continuously. Even after $y = y_1^{(u)}$ traverses this minimum point, $|\Delta_m|$ and $-W_m$ still increase continuously.

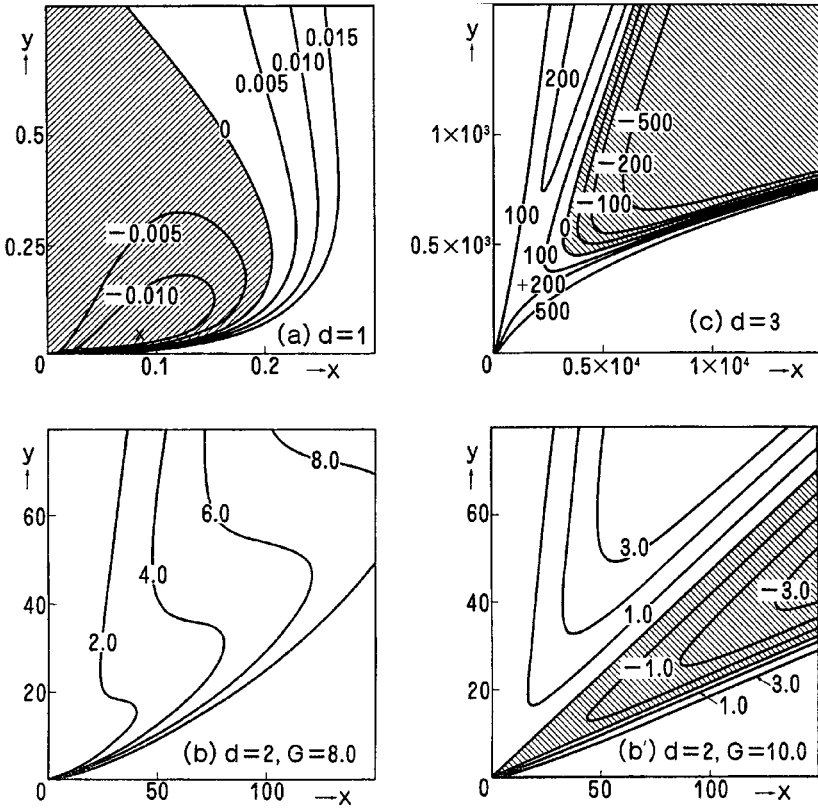


Fig. 9.10 (a)–(c) Adiabatic potentials for an electron in an elastic medium. The values of x and $y^{-1/2}$ are proportional to the strength of distortion and the radius of the distorted region, respectively. The energy is zero along the ordinate and negative along the hatched region. (b) and (b') represent the weak and strong coupling cases, respectively, of the two-dimensional lattice.³⁶

In contrast, the stable state in a three-dimensional system changes discontinuously from the free state $F(\Delta = 0)$ to the self-trapped state $S[\Delta \neq 0, a = (\gamma_3)^{1/3}a_0]$ being the minimum possible value] when g exceeds $g_c = 2.3$ (namely, when $y = y_3^{(u)}$ reaches the negative energy region in Fig. 9.10(c)). The height of the potential barrier from F- to S-state (the saddle point marked with + in Fig. 9.10(c) is given by $\varepsilon_b = 0.79g^{-2}B$. It should also be noted that the F-state remains metastable for $g_c < g < \infty$ and that the S-state is metastable for $g'_c < g < g_c$ where $g'_c = 1.86$ corresponds to $y_3^{(u)} = y$ (saddle point). The result justifies a posteriori the Ansatz $\phi_k(\mathbf{R}_n) = \delta_{n0}$ for the trial wave function which led to the discontinuous transition from F- to S-states in Section 9.4.

The situation is varied in a two-dimensional lattice: as $g = (\pi/16)G$ exceeds $g_c = 1.83$, the negative energy region sets in like an opening fan (Fig. 9.10(b')).

The stable state changes discontinuously from F- to S-state at $g = g_c$; moreover, there is neither a potential barrier nor the F–S coexistence range of g_c , in contrast to the three-dimensional system. It should be noted, however, that this peculiar feature of the two-dimensional system is lost as soon as the band energy deviates from the parabolic form.

The above results for one- to three-dimensional cases are in agreement with those obtained before by the variational method. One should be more careful in ascribing these features to the dimensionality effect only. In fact, the non-parabolicity of the electronic band $\varepsilon_c(\mathbf{k})$ completely changes the peculiar F–S transition in the critical two-dimensional lattice, and can significantly influence the gradual F–S change in a one-dimensional lattice.³⁷

It is of interest to see how the situation depends on anisotropy. Pertsch and Rössler³⁸ studied the effect of anisotropy in a three-dimensional system, with particular attention to the limiting situations of nearly one- and nearly two-dimensional systems.

So far we have considered the multistability of the electron–phonon system, which of course also applies to the hole–phonon system. The situation is more varied in the exciton–phonon system if we consider the phonon effect on the electron–hole *relative* motion, in addition to the translational motion. With the use of various forms of trial wave function for an electron and a hole *mutually* attracting through the Coulomb force and *individually* interacting with the acoustic phonon field through respective deformation potentials Ξ_e and Ξ_h , Sumi³⁹ obtained the phase diagram for the stable states, as shown in Fig. 9.11 for the case of $m_e : m_h = 1 : 4$. (Note that the lower half of the phase diagram is the inversion of the upper half with respect to the origin 0 since the simultaneous sign changes of Ξ_e and Ξ_h result in the same energy but with dilation being replaced by contraction.) In addition

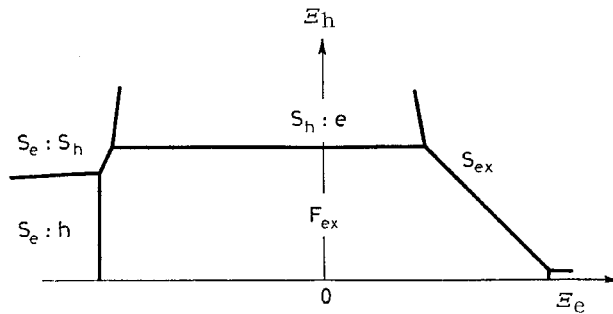


Fig. 9.11 Phase diagram for the stable state of an electron–hole (e–h) pair with Coulomb attraction, on the coordinate plane of electron and hole deformation potential constants. The e–h mass ratio is taken to be 1 : 4. Due to A. Sumi.³⁹

to the region F_{ex} where the free exciton state is stable, the region $S_h : e$ where the hole becomes self-trapped with the electron loosely bound by its Coulomb field and the similar but opposite region $S_e : h$, Sumi found two new phases. In the region S_{ex} , Ξ_e and Ξ_h act *constructively* to make electron and hole self-trapped on the same place under a *less imposing* condition than for individual self-trapping (as seen from the slanted borderline between F_{ex} and S_{ex} regions). In the region $S_e : S_h$, Ξ_e and Ξ_h act *destructively* to make electron and hole self-trapped on *different* but *neighboring* (due to Coulomb attraction) sites. The same and opposite signs of Ξ_e and Ξ_h result in attractive and repulsive phonon-mediated electron-hole interactions, respectively. In the opposite case, the electron wants to dilate, and the hole to contract, the surrounding lattice (or vice versa), being incompatible if they were on the same site.

We will now survey the results of experimental studies which are related to the foregoing arguments on the relative stability of free and self-trapped states, leaving the microscopic structure of the latter to a subsequent section. A few typical examples of electronic carriers – electrons, holes and excitons – are shown in Fig. 9.12, being assigned to appropriate regions of the phase diagram presented in Fig. 9.9 according to their types on the basis of relevant experimental facts. While all

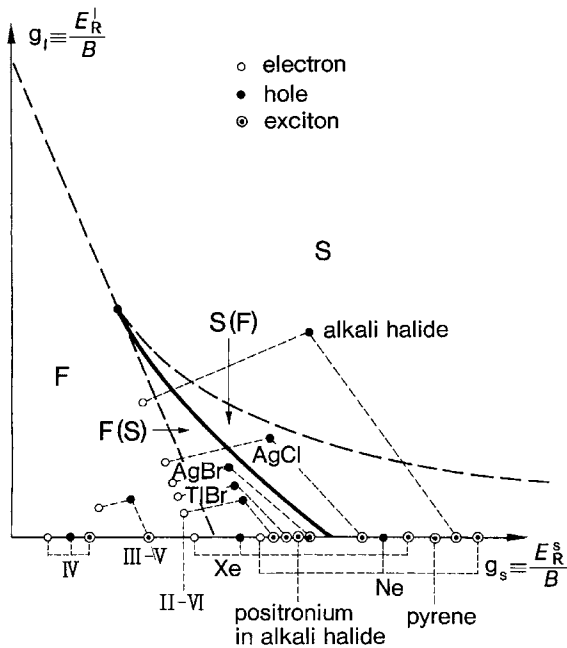


Fig. 9.12 Electrons, holes and excitons in typical insulators allocated in appropriate regions of the phase diagrams of Fig. 9.8 on the basis of various experimental studies.

electronic carriers in IVth column and III–V compound semiconductors belong to the F-region, the holes in many ionic crystals belong to S- or S(F)-regions in contrast to the electrons which are mostly F-type. Self-trapped holes²⁵ and self-trapped excitons⁴⁰ in alkali halides were recognized early on. Excitons, because of the absence of long-range coupling with phonons due to their electrical neutrality, are put on the g_s -axis which has no pure S region (see Fig. 9.12). This means that the optically-created excitons, initially in the F-state, have to overcome potential barriers with finite height ε_b before they relax to the S-state even when the latter is more stable: S(F)-region. This barrier has been observed in a number of materials in the study of coexistent emission bands from F- and S-excitons, with ε_b being estimated in alkali halides by Lusichik *et al.*,⁴¹ Hayashi *et al.*,⁴² Nishimura *et al.*⁴³ and Khiem and Nouhailhat⁴⁴ from the decrease in temperature of the intensity of the emission bands from the metastable F-excitons. In general, the emission band from the self-trapped exciton is Stokes-shifted from the absorption peak and has a large width due to the strong lattice relaxation in contrast to that from the unrelaxed free exciton which is a resonant sharp line, as will be described in Chapter 10. The temperature dependence of the intensity ratio of the two types of emission bands in pyrene and α -perylene, as studied by Matsui *et al.*,⁴⁵ indicates that the populations of the F(metastable)- and S(stable, excimer)-states are in quasi-thermal equilibrium (in contrast to the case of alkali halides), wherefrom they estimate the energy difference between the two states.

It is worth noting that the holes and excitons are self-trapped in AgCl but are free in AgBr in spite of the similarity of the two materials in the electronic structures of their ground and *unrelaxed* excited states. It indicates that the carriers (holes and excitons) in the two materials are to be located probably close together but definitely on different sides of the discontinuity line, as shown in Fig. 9.12. Even more direct evidence for the discontinuity line is provided by the observation by Kanzaki, Sakuragi and Sakamoto⁴⁶ of the abrupt change of the exciton emission bands of $\text{AgBr}_{1-x}\text{Cl}_x$ alloys from the F-type to the S-type as the composition x exceeds 0.45. The mobility measurements in AgBr by Hanson⁴⁷ and TlBr by Kawai *et al.*⁴⁸ indicate that the positive holes in these materials are to be allocated to the F(S) region: the anomalous decrease in their mobilities with increasing temperature can be explained only by assuming a metastable S-state a fraction of an electron-volt above the F-state.⁴⁹ According to Takahei and Kobayashi,⁵⁰ excitons in thallous halides are also near the marginal situation: although they are F-type both in TlBr and TlCl, they are subject to extrinsic (impurity-assisted) self-trapping in TlCl:Br; the impurity-bound excitons appear only in the luminescence (Stokes-shifted broad band, indicating a significant lattice relaxation) and not in the absorption spectra. Extrinsic self-trapping^{51,52} will be described in more detail in Section 13.7.

9.6 Microscopic structures of self-trapped electrons, holes and excitons

According to the foregoing section based on the continuum model of the lattice, the self-trapped electron, in contrast to the free electron, wants to take as small an extension as is compatible with its finite transferability to neighboring sites on the given structure of the lattice. For the study of the self-trapped state itself, therefore, it is of primary importance to consider its local microscopic structures, which however vary depending on the material being examined. Rather than attempting a systematic classification or theoretical deduction of this variety, we will first describe individual examples of self-trapped electrons, hole and excitons found experimentally in various materials and then explain them on the basis of (i) the atomic origins of the conduction, valence and exciton bands, (ii) the nature of bonds forming the crystal lattice and (iii) the types of electron–phonon interaction. See also the book by Song and Williams⁵³ for an extensive study of these problems based on the quantum-mechanical calculation to solve self-consistently the electronic structures and atomic rearrangements of the self-trapped excitons in various materials.

The simplest insulator consisting of a single element is the rare gas, in its gas, liquid and solid phases. Rare-gas atoms, due to their closed-shell structures, can exert only a weak van der Waals force upon each other, resulting in small cohesive energies and low melting points (He, due to its light mass and large zero-point motion, remains a liquid down to 0 K, and can be solidified only under applied pressure, being the only such element). It is for this reason that the gas, liquid and solid phases of the same rare-gas element have more or less similar absorption spectra, the corresponding lines being only slightly displaced among the different phases. The same applies to some extent even to the emission spectra, namely to the relaxed excited states,⁵⁴ as will be seen later.

Relaxed electrons in solid and liquid ^4He and in liquid Ne are considered to be self-trapped,⁵⁵ in view of their mobilities, as low as 10^{-4} – 10^{-2} $\text{cm}^2 \text{V}^{-1} \text{s}^{-1}$, the regime of the hopping motion. In fact, a *bubble*, a cavity in the medium, is formed around the electron, due to the negative electron affinity (measured to be -1.05 eV in liquid He, and -1.3 eV in Ne though in the solid state, referred to the vacuum level as zero) which makes the electron repel the surrounding medium. It corresponds to the negative deformation potential Ξ_c in our terminology which, if large enough, results in the self-trapped state with dilation of the surrounding lattice (see Section 9.4). The effective mass of this negative carrier in liquid He was experimentally obtained⁵⁶ as $m_-^* = 76M(^4\text{He})$ which corresponds to a 12 \AA radius of the cavity. The electronic structure and the cavity radius were discussed by Fowler and Dexter⁵⁷ in terms of competition between the kinetic energy of the confined electron and the surface tension of liquid He. The electrons in solid

Ne, and in solid and liquid Ar, Kr and Xe have high mobilities of the order of thousand $\text{cm}^2 \text{V}^{-1} \text{s}^{-1}$ and are considered to be in the nearly free state. The electrons in solid Ne are presumed to be particularly close to, but still below the critical condition for self-trapping in view of the situation in the liquid case mentioned before and also of the later statement on the formation of a bubble around an exciton.

In contrast, the positive hole in solid and liquid He II is considered to produce a cluster of He atoms, called a *snow-ball*,⁵⁸ with effective mass $m_+^* = 45M(^4\text{He})$ in the solid. It seems to be a self-trapped state due to a positive deformation potential resulting in contraction of the surrounding lattice. The large radii of the bubbles and the snow-balls are to be ascribed to the small cohesive energies (as seen from the low melting points) of He and Ne.

Positive holes in other rare-gas solids also have low mobilities, less than $10^{-2} \text{cm}^2 \text{V}^{-1} \text{s}^{-1}$, and are considered to be in the self-trapped state of a molecular type: R_2^+ (R denotes a rare-gas atom) in analogy to the self-trapped holes in alkali halides with a well-established molecular structure. The valence bands in solid rare gases (except He) originate from the highest occupied np -shells ($n = 2, 3, 4$ and 5 for Ne, Ar, Kr and Xe respectively) while the valence bands of alkali halides originate from the highest occupied np -shells of halide ions ($n = 2, 3, 4$ and 5 for F^- , Cl^- , Br^- and I^- respectively). The self-trapped hole was first found in an alkali halide by Känzig²⁵ who determined its structure, unambiguously, with the use of electron spin resonance experiments, to be of a molecular type: X_2^- (X denotes a halogen atom). The triply-degenerate atomic p -orbitals are oriented perpendicularly to each other, and tend to form a covalent bond with the p -orbital of opposite sign of a neighboring atom if both atoms come closer to each other than their normal distance in the lattice. To be more exact, in a closer pair of X^- , two electrons occupy the bonding orbital and two the antibonding orbital, the remaining eight being in orbitals perpendicularly oriented and irrelevant to our problem. The destabilized antibonding orbital tends to release one of the paired electrons by recombining with a positive hole, if one such is wandering about. Namely, a positive hole can be captured by a pair of X^- if the latter are sufficiently closer than the normal distance, thus forming a stable center: X_2^- . One could describe this simply as the result of a positive hole attracting a pair of X^- together, but what is important here is the *oriented* nature of *degenerate* p -orbitals, in contradistinction to the unoriented s -orbitals which in fact cannot form a molecular-type self-trapped state as far as the available experiments are concerned. In fact, the only example of a molecular-type self-trapped *electron* is in PbCl_2 and PbBr_2 ,^{59,60} whose conduction bands originate from the $6p$ -orbital of Pb (note that p -originated conduction bands are rather few among simple insulators).

The relaxed state of an exciton naturally reflects the natures of those of their constituents, the electron and the hole. The most intense emission bands in all

condensed rare gases are considered to originate from molecular-type self-trapped excitons, as seen from their resemblance to those of the excimers (excited dimers: X_2^*) in the gases exploited in “excimer lasers”. The bond formation is ascribed to the positive hole in the p-originated valence band in Ne, Ar, Kr and Xe, but rather to the resonant transfer of the $1s \rightarrow 2p$ excitation energy in the case of He (the same mechanism may partly be responsible to the formation of excimers in the above rare gases); the positive hole in He is 1s-originated while the excited electron is busy in forming the bubble around the already relaxed molecular-type exciton. In the case of condensed He, there appears also a weaker emission band with vibrational structures (characteristic of a molecule) closer to the absorption band, which corresponds to the initial stage of excimer formation. The condensed Ne has both the atomic (a)- and molecular (m)-type self-trapped excitons (STE). According to the time-resolved absorption spectroscopy of solid Ne under electron pulse excitation due to Suemoto and Kanzaki,⁶¹ the a-STE is initially formed, followed by the growth of a bubble around it whose rate is determined by the inward diffusion of thermal vacancies as is evidenced by the same activation energy. At a later stage (or if induced by laser irradiation), the m-STE is formed which also is followed by the growth of a bubble around it. It seems that the positive hole contributes to the formation of an excimer while the electron to the formation of bubbles around a- and m-STEs. It is the negative electron affinity of Ne that produces the a-STE as a relaxed excited state in addition to the m-STE.

The role of an electron in m-STE is different in ionic crystals, due to the Coulomb interaction between the electron and the ions. For instance, the contraction of the $X^- - X^-$ distance due to the self-trapped hole in an alkali halide causes a fractional anion vacancy at each end of the X_2^- center. So to speak, the net positive charge, $+e$, of the m-STH (self-trapped hole) is divided into two centers at these fractional vacancies. The electron captured by this two-centered Coulomb potential may cause instability of this atomic configuration against the displacement of X_2^- as a whole in either of two directions along its molecular axis which results in the formation of an entire anion vacancy.

Song and collaborators⁶² pointed out, on the basis of their theoretical calculations, that the most stable state of self-trapped excitons in many alkali halides is in such an *off-centered* configuration (with broken inversion symmetry) instead of the *on-center* configuration which had thus far been accepted. The off-center STE is, so to speak, the nearest pair of F- and H-centers, while more distant pairs correspond to what had been considered as the independent defects formed by irradiation. The symmetry-broken STE in alkali halides can also be considered as an example of the self-decomposed exciton, $S_e : S_h$, which had been predicted by A. Sumi;³⁹ as far as the anion sublattice is concerned, the positive hole and the electron want, respectively, to contract and dilate the surrounding lattice, namely through positive and

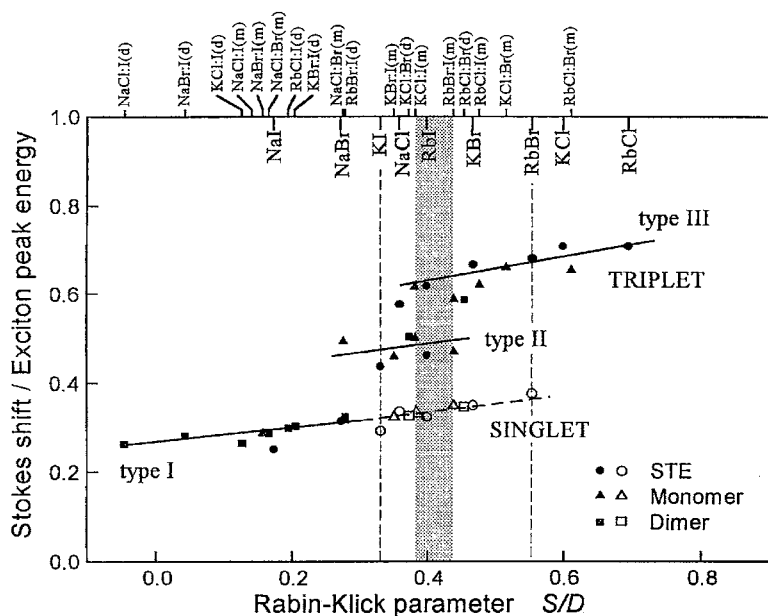


Fig. 9.13 The correlation between the Stokes shift of exciton luminescence (the ordinate) and the Rabin–Klick parameter (the abscissa) in various alkali halides, indicating three types of lattice relaxation in increasing order of magnitude. (1) STE, (2) monomer (m) and (3) dimer (d) mean an exciton self-trapped around a nearest pair of (1) host halogen atoms, (2) host and guest (impurity) atoms and (3) guest halogen atoms, respectively. Reproduced from Kan’no, Matsumoto, Kayanuma.⁶³

negative deformation potentials ($\Xi_h > 0$, $\Xi_e < 0$), which are compatibly realized when the electron and the hole are a little decomposed, as they are. Kan’no *et al.* showed,⁶³ on the basis of careful analysis of their time-resolved spectroscopy of emission spectra and by correlating the available data on the Stokes shift (a measure for the strength of lattice relaxation for self-trapping) with the Rabin–Klick parameter (the ratio of the gap distance between adjacent halide ions to the diameter of a halide ion) as shown in Fig. 9.13, that STEs in alkali halides can be unambiguously classified into three types: (I) on-center, (II) weakly off-center, and (III) off-center types. This tendency is qualitatively consistent with calculations by Song and collaborators.⁶² The long-standing confusion on the interpretation of the varied emission spectra of STEs has been clearly resolved by these two groups.

PbCl_2 and PbBr_2 belong to very few examples of both the electron and the positive hole being relaxed into an m-exciton, the electron at a pair of cations (the conduction band originates from the 6p-orbits of Pb) and the positive hole at a pair of anions (the valence band from the hybrid np -orbits of the halogen and 6s-orbit of Pb). Iwanaga, Watanabe and Hayashi⁶⁴ observed long-lived blue–green emission mainly excited under ionizing radiation and short-lived blue emission under light,

creating an exciton. The former seems to be the tunneling recombination radiation of separated pairs of STH and STEL (self-trapped electron), while the latter to be the recombination radiation of the adjacent pairs (a modified form of the self-decomposed exciton), though this assignment is as yet tentative. The reason why self-trapping of the p-*electron* appears *only in a heavy metal* (such as Pb) compounds will be mentioned in Section 13.1.

A different type of a-STH and a-STE were found in AgCl whose valence band consists of hybrid Cl 3p-orbit and Ag 4d-orbit. The ESR study of STH shows its tetragonally distorted lattice configuration, indicating that the STH consists mainly of an Ag 4d-orbit.⁶⁵ The STE is considered to consist of an electron weakly bound by this STH. The absence of STH and STE in AgBr which has a similar electronic structure provides evidence of the discontinuity line in the phase diagram of Fig. 9.8 caused by short-range hole-phonon and exciton-phonon interactions.

Optical spectra of exciton in the phonon field

Overview

In this chapter, we will be concerned with one of the principal subjects of this book, the spectroscopic study of microscopic dynamical processes in matter, with the exciton in the phonon field as a model system. The linear response of matter to an electromagnetic wave of definite frequency reveals a component, with the same frequency, of the motion of the charged particles (electrons and nuclei) in the matter. The linearity of this response is usually assured under not-too-intense light due to the weak radiation–matter interaction. Therefore, the frequency dependence of the linear response such as the lineshape of absorption spectra (imaginary part of susceptibility apart from an unimportant factor), tells us what is going on in the microscopic world.

Both light and matter have the dual nature of wave and particle and obey the uncertainty principle of quantum mechanics, facts which govern all aspects of the spectroscopic study of matter. An exciton, a typical elementary excitation in an insulating solid, behaves as a quasi-particle with definite dispersion (energy–momentum relation), and can be created by annihilation of an incoming photon. Due to the energy–momentum conservation rule, this elementary process can take place only with a photon with definite energy equal to that of the exciton with the same wave vector; namely, the absorption spectrum consists of an infinitely sharp line. Monochromatic light with infinite duration time corresponds to an exciton with definite energy and hence infinite lifetime. However, a real exciton, which is under a fluctuating potential field due to lattice vibrations, has a *finite lifetime* τ of *being scattered* to states with different wave vectors, and the resulting energy width \hbar/τ of the exciton is reflected in the *lifetime broadened Lorentzian* spectrum. Namely, the shape and width of the absorption spectrum tells us about the exciton scattering by lattice vibrations (Section 10.2).

If the amplitude of the local potential fluctuation is larger than the intersite transfer energy, one should consider a localized exciton subject to occasional transfer to a neighboring site, rather than a free exciton subject to occasional scattering. In this situation, the *fluctuation of local potential* obeying a Gaussian distribution is directly reflected in a *Gaussian lineshape* of the absorption spectrum. The two contrasting pictures, the *localized* and *delocalized* (free) excitons, are on the opposite extremes of the *position–momentum uncertainty*. In fact, the gradual transition from a localized to a delocalized exciton as reflected in the spectral shape and width, known as *motional narrowing*, takes place as the intersite transfer energy exceeds the local fluctuation energy, thus *preventing the exciton from experiencing the latter to its full extent* in accordance with the uncertainty principle (Section 10.3).

In the low-energy tail of the absorption spectrum away from its peak, however, the *momentarily localized states* of an exciton under temporal and spatial fluctuations show up even in the case of weak exciton–phonon scattering (Section 10.6). The duration time of this momentary localization is of the order of (or less than) the reciprocal of energy distance between the peak and the tail region we are concerned with, according to the uncertainty principle.

In this way, the spectral shape of an exciton line is, so to speak, a window or a slit through which one can look into the microscopic world. In the case of multiple exciton lines corresponding to different internal states, or of multiple slits in our analogy to quantum mechanics, the *interference of beams* from different slits – *from different internal states as intermediate states* – shows up in the spectral anomaly (Section 10.2), known as the Fano effect in atomic spectroscopy. The matrix elements of the exciton–phonon interaction between different internal states, which are responsible for this interference effect, are also relevant to the polaron effect of the electron–hole relative motion (Section 10.1).

The quantal aspect of the fluctuation due to lattice vibrations shows up as phonon structures in the spectra (Sections 10.4 and 10.5), in a way which depends upon the situation. In the motionally-narrowed situation, the phonon structures are implicitly contained in the energy dependence of the *width function* in the Lorentzian expression of lineshape, while in the opposite situation they show up more directly as multiple phonon sidebands through the exciton–phonon *coupling function*.

While the absorption spectra tell us about the behaviors of an exciton *immediately after its creation*, the emission spectra tell us about the behavior *immediately before its annihilation*. This means also that one can *speculate about the exciton behavior between its birth and death* by studying the *correlation of absorption and emission spectra* (Section 10.6). Since the radiative lifetime of an exciton is finite, the two spectra are not entirely independent of each other. In fact, with high enough sensitivity and resolving power one can in principle find, in the absorption spectra,

the faint zero-phonon line due to the self-trapped state which can be reached only after quantum tunneling from the free state in real time. This does not contradict the uncertainty principle since one is concerned here with an extremely sharp zero-phonon line whose recognition needs extremely long time, both being limited by the radiative lifetime of the self-trapped exciton. In the case of no self-trapping, the emission spectrum of a direct exciton is to be described in terms of polaritons which are populated around the bottleneck (Section 10.8) in a way *different from thermal equilibrium* due to their finite lifetime.

Readers, after having finished this chapter, are recommended to come back here to confirm or to reinterpret, in their own ways, what has been stated in this overview. A more explicit pursuit of the exciton dynamics throughout its entire life span can be performed by time-resolved spectroscopy or pump-and-probe type experiment, as will be described in later chapters.

10.1 Exciton–phonon interaction and the polaron effect of an exciton

While the exciton is an elementary excitation in a many-electron system, each electron ($i = 1, 2, \dots$) is subject to the phonon interaction $H_I(\mathbf{r}_i)$, as given by (9.1.17) and (9.4.3) in the cases of optical and acoustic mode, respectively. Confining ourselves to a single-exciton subspace, let us calculate the matrix element of the scattering of an exciton between different states $\Psi_{\lambda K}$. Noting that they are a linear combination of the type (8.3.2), and that the matrix element between the one-electron excited states $\Psi_{v\mu k K}$ on the r.h.s. of (8.3.2) is given as the sum of those of electron–phonon and hole–phonon interactions with respective coupling coefficients $V_c(\mathbf{q})$ and $-V_v(\mathbf{q})$, one obtains^{1,2}

$$\begin{aligned} & \int \int \cdots d\mathbf{r}_1 d\mathbf{r}_2 \cdots \Psi_{\lambda K}^* \left[\sum_{i=1}^N H_I(\mathbf{r}_i) \right] \Psi_{\lambda' K'} \\ &= V_{\lambda\lambda'}(\mathbf{K} - \mathbf{K}') b_{\mathbf{K}-\mathbf{K}'} + V_{\lambda\lambda'}^*(-\mathbf{K} + \mathbf{K}') b_{-\mathbf{K}+\mathbf{K}'}^\dagger, \end{aligned} \quad (10.1.1)$$

where

$$V_{\lambda\lambda'}(\mathbf{q}) \equiv V_c(\mathbf{q})\eta_{\lambda\lambda'}(p_e\mathbf{q}) - V_v(\mathbf{q})\eta_{\lambda\lambda'}(-p_h\mathbf{q}), \quad (10.1.2)$$

$$\eta_{\lambda\lambda'}(\mathbf{q}) \equiv \int d\mathbf{R} F_\lambda'(\mathbf{R}) F_{\lambda'}'(\mathbf{R}) \exp(i\mathbf{q} \cdot \mathbf{R}), \quad (10.1.3)$$

$$p_{e,h} \equiv m_{h,e}/(m_e + m_h). \quad (10.1.4)$$

The form factors $\pm\eta_{\lambda\lambda'}(\pm p_{e,h}\mathbf{q})$ represents the effective charges of the electron and the hole in the exciton seen by the phonon \mathbf{q} when the exciton is scattered from λ' to λ of the internal motion. Note that $\pm p_{e,h}\mathbf{R}$ represent the coordinates of the electron and the hole, respectively, relative to their center-of-mass coordinate.

For the *intraband* (here the band means an *exciton* band) scattering of a 1s exciton ($\lambda = \lambda' = 1s$), as an example, one obtains

$$\eta(q) = [1 + (a_B q/2)^2]^{-2} \quad (10.1.5)$$

since $F_{1s}(\mathbf{R}) \propto \exp(-R/a_B)$. This means that those phonons with wavelengths $2\pi/q$ larger than the exciton radius a_B are ineffective for scattering the exciton due to the cancellation effect among different phases of the wave. In the case of optical phonon interaction for which $V_c(\mathbf{q}) = V_v(\mathbf{q})$ (electrostatic potential being common), we have moreover the cancellation effect at long wavelengths of the intraband scattering due to the charge neutrality of an exciton, as seen from

$$V_{1s,1s}(\mathbf{q}) = V_c(\mathbf{q})\{[1 + (p_e a_B q/2)^2]^{-2} - [1 + (p_h a_B q/2)^2]^{-2}\}. \quad (10.1.6)$$

Since $V_{1s,1s}(\mathbf{q}) \propto q$ in contrast to $V_c(\mathbf{q}) \propto q^{-1}$ for small q , the exciton has no long-range interaction with optical phonons. This is the reason why the exciton always (a single electron or hole not always!) has to surmount a potential barrier of finite height before it is relaxed into a self-trapped state even if the latter is more stable, as mentioned in Sections 9.5 and 9.6 and in particular in connection with Fig. 9.12. A more elaborate study of this problem is presented in Ref. [3], where the coupling coefficient (10.1.6) for a free exciton should be used to study how the potential ascends (rises) towards the barrier although the situation changes completely as soon as self-localization starts.

While the intraband scattering matrix elements are subject to the partial (complete in the particular case of $p_e = p_h$) cancellation of the electron and hole charge distributions, the interband elements cause mixing among different internal states, resulting in a change in the relative motion. Haken⁴ started from a pair of an electron polaron and a hole polaron instead of a bare electron–hole pair as mentioned above, and described this problem in terms of the effective potential between the electron polaron and a hole polaron, which turned out to be given by $-e^2/4\pi\epsilon_e r$ and $-e^2/4\pi\epsilon_s r$ at an electron–hole distance r small and large, respectively, compared to the sum of the polaron radii.

Extensive calculations of the renormalized exciton states were performed by Matsuura and Büttner,⁵ with the use of a variational method first proposed by Pollmann and Büttner.⁶ The point of this method consists in the determination of the electron–hole relative motion *self-consistently* with the renormalization of the translational motion by a Lee–Pines-type canonical transformation (see Section 9.2). The important results of the calculation are as follows. The polaron effect is insignificant for excitons with binding energy much larger than the optical phonon energy and hence with radius a_B much smaller than the sum of the polaron radii due to the cancellation effect in conformity with (10.1.6). The polaron effect is significant for excitons with larger radius *and* with much different electron and hole

effective masses. The consequence of their study on the optical phonon sideband will be mentioned in Section 10.4.

In the case of the acoustic mode interaction, the deformation potential constants Ξ_e and Ξ_h are independent quantities. The cooperative or destructive cases appear according as they have the same or opposite sign, instead of the cancellation effect mentioned above, as was also mentioned in connection with self-trapping in Section 9.6.

10.2 Phonon-induced interference in multicomponent exciton spectra

The effect of phonons on the exciton spectra have various aspects which have been studied more or less independently. The first is the temperature-dependent shifts and widths of exciton lines² – the *direct* exciton transitions represented by vertical arrows ($g \rightarrow \lambda' \mathbf{0}$) in Fig. 10.1 which is the energy scheme for one-electron excitations in an insulator. The second is the phonon-assisted (shown by wavy arrows) *indirect* exciton transitions⁷ ($g \rightarrow \lambda' \mathbf{0} \rightarrow \lambda \mathbf{K}$ in Fig. 10.1) contributing to the continuous spectra (because of continuous exciton band $E_{\lambda \mathbf{K}}$ in the final states) which form the tail parts of the broadened exciton lines. The third is the interference

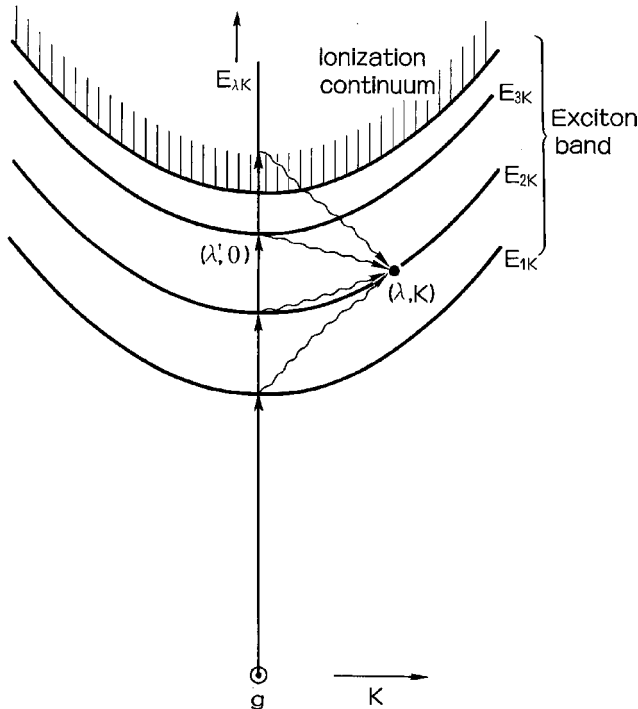


Fig. 10.1 Multichannel indirect transition consisting of optical creation (straight arrows) of an exciton with simultaneous emission or absorption of a phonon (wavy arrows).⁸

effect⁸ between different exciton lines caused by the fact that there are different intermediate states $\lambda' \mathbf{0}$ to reach the same final state $\lambda \mathbf{K}$, as shown by different sets of vertical followed by wavy arrows. The fourth is the phonon-originated fine structures (such as steps and even distinct peaks) in the exciton spectra, which are concealed in the energy dependence of the width function $\Gamma(E)$.⁹ All of these aspects can be incorporated into a unified theory as seen in the following heuristic derivation.

Let us start with the second-order perturbation theory for the indirect exciton spectra. Apart from unimportant factors, the rate of this second-order process – through various intermediate exciton states $\lambda' \mathbf{0}$ by virtual absorption of a photon with energy E , to the final state with the exciton being scattered to $\lambda \mathbf{K}$ by simultaneous emission or absorption of a phonon with wave vector $+\mathbf{K}$ – is given by

$$F(E) = \sum_{\lambda \mathbf{K} s \pm} \left| \sum_{\lambda'} (H_1)_{\lambda \mathbf{K} \lambda' \mathbf{0}}^{(s \pm)} P_{\lambda' g} / (E - E_{\lambda' \mathbf{0}}) \right|^2 \times \delta(E_{\lambda \mathbf{K}} \pm \hbar \omega_{s, \mp \mathbf{K}} - E). \quad (10.2.1)$$

Here $P_{\lambda' g}$ is the matrix element of the polarization density for the optical transition ($g \rightarrow \lambda' \mathbf{0}$) and

$$(H_1)_{\lambda \mathbf{K} \lambda' \mathbf{0}}^{(s \pm)} = \begin{cases} V_{\lambda \lambda'}^{(s)*}(-\mathbf{K})(n_{s, -\mathbf{K}} + 1)^{1/2} \\ V_{\lambda \lambda'}^{(s)}(\mathbf{K})(n_{s, \mathbf{K}})^{1/2} \end{cases} \quad (10.2.2)$$

represents the matrix element for exciton scattering $\lambda' \mathbf{0} \rightarrow \lambda \mathbf{K}$ by emission (+) or absorption (−) of a phonon of the s th branch. The thermal average over the phonon numbers, $n_{s, \mathbf{K}}$, is to be taken on the r.h.s. of (10.2.2).

It is convenient to introduce the second-order self-energy *matrix* $\Sigma_{\mathbf{K}'}(E)$ defined by

$$\begin{aligned} \Sigma_{\lambda \lambda' \mathbf{K}'}(E) &\equiv \Delta_{\lambda \lambda' \mathbf{K}'}(E) + i \Gamma_{\lambda \lambda' \mathbf{K}'}(E) \\ &= \sum_{\lambda \mathbf{K} s \pm} (H_1)_{\lambda' \mathbf{K}'}^{(s \mp)} (H_1)_{\lambda \mathbf{K} \lambda' \mathbf{0}}^{(s \pm)} \\ &\quad \times \{ P[E - E_{\lambda \mathbf{K}} + \hbar \omega_{s, \mp(\mathbf{K}-\mathbf{K}')}]^{-1} \\ &\quad + i \pi \delta[E - E_{\lambda \mathbf{K}} + \hbar \omega_{s, \mp(\mathbf{K}-\mathbf{K}')}] \}. \end{aligned} \quad (10.2.3)$$

The real and imaginary parts of a *diagonal* element ($\lambda' = \lambda$) of this matrix are the self-energy and the half-width, respectively, of the $\lambda' \mathbf{K}'$ exciton in the lowest order of the perturbation theory. It is easy to see that the matrices Δ and Γ are Hermitian, with Γ being non-negative. Similarly we define a diagonal matrix $H_{e\mathbf{K}}$ with its $(\lambda \lambda)$ element being given by the exciton energy $E_{\lambda \mathbf{K}}$, and a one-column matrix P with its λ th row element given by $P_{\lambda g}$ and its adjoint one-row matrix P^\dagger . We can then

rewrite (10.2.1) more compactly, with the use of the matrix representation, as

$$F(E) = \pi^{-1} \sum_{\lambda' \lambda''} P_{g\lambda'}^\dagger (E - E_{\lambda' \mathbf{0}})^{-1} \Gamma_{\lambda' \lambda''} \mathbf{0}(E) (E - E_{\lambda'' \mathbf{0}})^{-1} P_{g\lambda''} \quad (10.2.4a)$$

$$= \pi^{-1} P^\dagger (E - H_{e0})^{-1} \Gamma_0(E) (E - H_{e0})^{-1} P. \quad (10.2.4b)$$

Decomposing the cross terms ($\lambda' \neq \lambda''$) in (10.2.4a) as

$$\begin{aligned} & (E - E_{\lambda' \mathbf{0}})^{-1} (E - E_{\lambda'' \mathbf{0}})^{-1} \\ &= (E_{\lambda' \mathbf{0}} - E_{\lambda'' \mathbf{0}})^{-1} [(E - E_{\lambda' \mathbf{0}})^{-1} - (E - E_{\lambda'' \mathbf{0}})^{-1}], \end{aligned} \quad (10.2.5)$$

we can rearrange the double summation in it into a single summation:

$$F(E) = \pi^{-1} \sum_{\lambda} f_{\lambda} [\Gamma_{\lambda \lambda \mathbf{0}}(E) + A_{\lambda}(E)(E - E_{\lambda \mathbf{0}})] / (E - E_{\lambda \mathbf{0}})^2 \quad (10.2.6)$$

where

$$f_{\lambda} \equiv |P_{g\lambda}|^2, \quad (10.2.7)$$

$$f_{\lambda} A_{\lambda}(E) \equiv \sum_{\lambda' (\neq \lambda)} 2\text{Re}\{P_{g\lambda} \Gamma_{\lambda \lambda' \mathbf{0}} P_{g\lambda'}\} / (E_{\lambda \mathbf{0}} - E_{\lambda' \mathbf{0}}). \quad (10.2.8)$$

The divergence of each term in (10.2.6) at $E = E_{\lambda \mathbf{0}}$ is obviously an artifact of the second-order perturbation theory. Since the intermediate state $\lambda \mathbf{0}$ is subject to shift and broadening as given by (10.2.3), the denominator in (10.2.6) would have to be replaced by

$$[E - E_{\lambda \mathbf{0}} - \Delta_{\lambda \lambda \mathbf{0}}(E)]^2 + [\Gamma_{\lambda \lambda \mathbf{0}}(E)]^2$$

in higher-order theory. Then we obtain the *asymmetric* Lorentzian lineshape, as was pointed out in Ref. [2]. For a *complete* formal consistency, however, the diagonal and non-diagonal elements of the self-energy should together be incorporated into (10.2.4a) in the matrix form:

$$\begin{aligned} F(E) &= \pi^{-1} P^\dagger [E - H_{e0} - \Delta_0(E) - i\Gamma_0(E)]^{-1} \Gamma_0(E) [E - H_{e0} - \Delta_0(E) \\ &\quad + i\Gamma_0(E)]^{-1} P = \pi^{-1} P^\dagger \{[E - H_{e0} - \Delta_0(E) - i\Gamma_0(E)]^{-1} - [\text{h.c.}]\} P. \end{aligned} \quad (10.2.9)$$

In fact, it can be shown that (10.2.9) is the exact formula for the lineshape of multi-component exciton spectra provided that the following exact expression for the self-energy is used:

$$\begin{aligned} \Sigma_{K'}(E) &= \sum_{K s \pm} (H_I)_{K' K}^{s \mp} [E - H_{eK} \mp \hbar\omega_{s, \mp}(K - K')] \\ &\quad - \Sigma_K(E \mp \hbar\omega_{s, \mp}(K - K'))]^{-1} (H_I)_{K K'}^{s \pm} + (\text{higher-order terms}). \end{aligned} \quad (10.2.10)$$

Note that the matrix form of H_I with respect to $(\lambda \lambda')$ has been used in (10.2.10). See Ref. [8] for an exact derivation of (10.2.9).

Let us diagonalize the renormalized energy matrix by a transformation $T_K(E)$ (not unitary),

$$T_K(E)[H_{eK} + \Delta_K(E) + i\Gamma_K(E)]T_K(E)^{-1} = \bar{H}_K(E) + i\bar{\Gamma}_K(E). \quad (10.2.11)$$

Here $\bar{H}_K(E)$ and $\bar{\Gamma}_K(E)$ are real, diagonal matrices whose λ th element will be denoted by $\bar{E}_{\lambda K}(E)$ and $\bar{\Gamma}_{\lambda K}(E)$, respectively. Using (10.2.11), we can rewrite (10.2.9) as

$$\begin{aligned} F(E) &= (2\pi i)^{-1} \sum_{\lambda} [(P^\dagger T_0^{-1})_{\lambda} (T_0 P)_{\lambda} (E - E_{\lambda 0} - i\Gamma_{\lambda 0})^{-1} - (c.c.)] \\ &= \sum_{\lambda} \bar{f}_{\lambda} \pi^{-1} [\bar{\Gamma}_{\lambda 0} + \bar{A}_{\lambda} (E - \bar{E}_{\lambda 0})] / [(E - \bar{E}_{\lambda 0})^2 + \bar{\Gamma}_{\lambda 0}^2] \end{aligned} \quad (10.2.12)$$

where

$$(P^\dagger T_0^{-1})_{\lambda} (T_0 P)_{\lambda} \equiv \bar{f}_{\lambda} + i\bar{f}_{\lambda} \bar{A}_{\lambda}. \quad (10.2.13)$$

Note that the renormalized quantities with overbars are E dependent. If we neglect their energy dependence, each term in (10.2.12) represents an asymmetric Lorentzian lineshape, with a dip on the low- or high-energy side of the peak according as \bar{A}_{λ} , the degree of asymmetry, is positive or negative. As is obvious from (10.2.7), the expression for A_{λ} in the lowest-order approximation which resulted from the decomposition of the cross terms (10.2.5), the asymmetry results from the interference between different intermediate states $E_{\lambda 0}$. It is well known in atomic physics as the Fano effect, which takes place when a discrete spectrum overlaps with continuous spectra of different origin (different intermediate states in our case).

A beautiful example of the asymmetric lineshapes in multicomponent exciton spectra was observed in the yellow series excitons in Cu_2O as already shown in Fig. 8.3. In this case of forbidden band edge (case (b) of Section 8.4), only np excitons ($n = 2, 3, \dots$ for discrete states, and $n = k$ for ionization continuum) are optically allowed contributing to the λ th Lorentzian peaks in (10.2.12), while the phonons bring these direct excitons to entire exciton states (with $\mathbf{K} \neq 0$) including $\lambda = 1s$. The latter contribute to the continuous background of the absorption spectrum (nothing other than the Lorentzian tails) which starts at the $1s$ band edge (\pm phonon energies) reflecting its density of states. The Fano effect of a discrete line np takes place as its interference with this continuous background absorption although the latter comes also from various λ' -terms of the former. In other words, the Fano effect in this case takes place due to the interference between a direct transition line and the indirect transition continuum.

If the bottom of the lowest exciton band λ is not at $\mathbf{K} = 0$, but at \mathbf{K}_i (in general, there is more than one point with the same symmetry in the first Brillouin zone), the

absorption spectra at $T = 0$ starts as $[E - (\bar{E}_{\mathbf{K}_i} + \hbar\omega_{s,-\mathbf{K}_i})]^{1/2}$ for each phonon mode s reflecting the density of exciton states which rises as $[E - \bar{E}_{\mathbf{K}_i}]^{1/2}$ (see (10.2.3) for $\Gamma(E)$). At finite temperature, the absorption edge consists of several such steps with phonon absorption as well as emission,⁷

$$F(E) = \sum_s C_s \{n_{s,\mathbf{K}_i} [E - (\bar{E}_{\mathbf{K}_i} - \hbar\omega_{s,\mathbf{K}_i})]^{1/2} + (n_{s,\mathbf{K}_i} + 1) [E - (\bar{E}_{\mathbf{K}_i} + \hbar\omega_{s,\mathbf{K}_i})]^{1/2}\} \quad (10.2.14)$$

where $[\dots]^{1/2}$ is to be taken as zero when the argument is negative. If symmetry forbids the scattering of any of the optically-allowed excitons $\lambda' 0 (M_{\lambda'g} \neq 0)$ to $\lambda \mathbf{K}_i$ by the (s, \mathbf{K}_i) phonon, the exponent of the s th absorption edge is three-halves instead of one-half since the $(\lambda' 0 \rightarrow \lambda \mathbf{K})$ scattering matrix element increases proportionally to $|\mathbf{K} - \mathbf{K}_i|$.

The direct absorption edge of the form (10.2.14) has been observed in a variety of crystals such as Si, Ge⁷ and silver halides.¹⁰ In thallous halides with the exciton-band bottom at the corner point (R) of the first Brillouin cube, all phonon modes are forbidden due to the high symmetry of the R point, and all edges rise with three-halves power of energy.¹¹

Other types of phonon structures concealed in the energy dependence of $\Gamma(E)$ will be described in Section 10.4.

10.3 Motional narrowing of exciton linewidth

We are now in a position to study the linewidth W (to be defined as *full* width at half maximum, hereafter) of the exciton λ which according to the general formula (10.2.12) is given by $2\bar{\Gamma}_{\lambda 0}(\bar{E}_{\lambda 0})$ (\hbar times the reciprocal scattering time $\tau_{\lambda 0}$ of the optically-produced exciton $\lambda 0$ by phonons of various modes), provided (i) the energy dependences of the renormalized quantities (the ones with overbars, inclusive of $\bar{\Gamma}_{\lambda 0}(E)$ itself) are negligible and (ii) the degree of asymmetry, \bar{A}_λ , vanishes. We will for the moment confine ourselves to the lowest exciton line for which condition (ii) holds since the line is not subject to superposition of background absorption due to higher exciton bands. Condition (i) holds approximately, for weak exciton-phonon coupling, if $E_{\lambda 0}$ is neither at the bottom nor at the top of the exciton band $E_{\lambda \mathbf{K}}$ since the state density $\rho_\lambda(E) \equiv \sum_{\mathbf{K}} \delta(E - \bar{E}_{\lambda \mathbf{K}})$ can be approximated by a constant within the range $\pm \Gamma_{\lambda 0}$ of $E = \bar{E}_{\lambda 0}$ if a possible anomaly due to the van Hove singularity is negligible. One must be more careful in the bottom (top) case since the energy dependence of $\bar{\Gamma}_{\lambda 0}(E)$ plays a crucial role for the lineshape. Neglecting the phonon energies for the moment, one obtains the width function from (10.2.2, 10.2.3)

$$2\Gamma_{\lambda 0}(E) = \hbar/\tau_{\lambda 0} = 2\pi \sum_{\mathbf{K},s} [|V^{(s)}(\mathbf{K})|^2 (2n_{s,\mathbf{K}} + 1)] \delta(E - E_{\lambda \mathbf{K}}). \quad (10.3.1)$$

If the quantity $[\dots]$ is independent of \mathbf{K} as is the case for long-wavelength acoustic phonon scattering of an exciton near the band bottom (note that the form factor η defined by (10.1.3) can be approximated by unity, $|V^{(s)}(\mathbf{K})|^2 \propto K$ (see (9.4.3)) and $(2n_{s,\mathbf{K}} + 1) \simeq 2k_B T / \hbar c_\ell K$), one can rewrite (10.3.1) as

$$2\Gamma_{\lambda 0}(E) = 2\pi \left[N \sum_s |V^{(s)}(\mathbf{K})|^2 (2n_{s,\mathbf{K}} + 1) \right] N^{-1} \rho_\lambda(E). \quad (10.3.2)$$

For qualitative purposes, one can replace the factor N in $[\dots]$ by summation over all \mathbf{K} . The sum can be written as D^2 similarly to (4.3.12). In fact, D^2 represents the mean squared amplitude of the local potential fluctuation to which the exciton is subject. One can thus write the width function, irrespective of the position of $E_{\lambda 0}$ within the band $[E_{\lambda K}]$, as

$$2\Gamma_{\lambda 0}(E) \sim 2\pi D^2 N^{-1} \rho_\lambda(E). \quad (10.3.3)$$

If $E_{\lambda 0}$ is neither at the bottom nor at the top of the band λ , one can replace the normalized density of states $N^{-1} \rho(E_{\lambda 0})$ by its average $1/2B$ for an order of magnitude estimation, where $2B$ is the width of the exciton band. The resulting width is estimated as

$$2\Gamma_{\lambda 0} \sim (\pi/2) D \times (D/B). \quad (10.3.4)$$

This means that the linewidth $W = 2(2 \ln 2)^{1/2} D$ for the localized excitation (see (4.3.10) to (4.3.12)) is reduced by a factor $\sim (D/B)$ due to the translational motion of excitation for a wide-band exciton: $B \gg D$ (*weak scattering* condition as it is called).¹ This is an example of the *motional-narrowing* effect which is well known in the visible to microwave spectroscopy of moving entities.^{12,13} In our case, the exciton is transferred to a neighboring site within a time $\sim \hbar/B$, which is shorter than the minimum time \hbar/D required by the uncertainty principle for that exciton to fully experience the local fluctuation D . The width is proportional to T at high temperatures due to (4.3.12), in contrast to $T^{1/2}$ for a localized excitation.

If $E_{\lambda 0}$ is at the bottom of the exciton band as was often assumed in Chapter 8, the situation is more complicated since the energy dependence of $\Gamma_{\lambda 0}(E) (\propto \rho(E) \propto (E - E_{\lambda 0})^{1/2})$ above, and $= 0$ below, the band bottom $E = E_{\lambda 0}$ plays a crucial role. Qualitatively, one can determine the width self-consistently by putting

$$E - E_{\lambda 0} = \Gamma_{\lambda 0}(E). \quad (10.3.5)$$

As it implies, one has to take the value of the width function $\Gamma_{\lambda 0}(E)$ at the peak energy E of the absorption spectrum which is presumably a half-width $\Gamma_{\lambda 0}$ above the band bottom $E_{\lambda 0}$. Assuming a parabolic exciton band which ends at K_0 with total bandwidth $2B = \hbar^2 K_0^2 / 2m_{\text{ex}}$, one can write the normalized density

of states as

$$N^{-1}\rho_\lambda(E) = (3/4\sqrt{2})B^{-3/2}(E - E_{\lambda 0})^{1/2}. \quad (10.3.6)$$

Putting this into (10.3.5), one obtains the linewidth:

$$W_\lambda \sim 2\Gamma_{\lambda 0}(E) \sim (9\pi^2/16)D \times (D/B)^3 = (9\pi^2/4)(E_R^{(ac)})^2 B^{-3}(k_B T)^2, \quad (10.3.7)$$

$$\text{where } E_R^{(ac)} \equiv (\Xi_e + \Xi_h)^2/2Ca_0^3 \quad (10.3.8)$$

(see (4.3.12) and (9.5.6)). Namely, the motional-narrowing effect is much more significant than in (10.3.4).¹⁴ The T^2 dependence of the linewidth W_λ indicates that two-phonon scattering (instead of one-phonon) makes a dominant contribution to the width. In fact, eq. (10.3.5) indicates that a typical exciton with wave vector \mathbf{K} contributing to the width satisfies

$$\tau_{\lambda 0}(\mathbf{K}) \sim \tau_c(\mathbf{K}) \quad (10.3.9)$$

where $\tau_c(\mathbf{K})$ is the duration time of an exciton-phonon collision given by

$$\tau_c(\mathbf{K}) = (2\pi/K)/(\hbar K/m_{ex}), \quad (10.3.10)$$

namely, the de Broglie wavelength of an exciton divided by its velocity. Equation (10.3.9) indicates that excitons with wave vector K or smaller (for which the equality \sim in (10.3.9) is to be replaced by the inequality $<$ since $\tau_{\lambda 0}(\mathbf{K}) \propto K^{-1}$) are subject to successive collisions inseparably. In other words, simultaneous two- and multiphonon scatterings predominate over one-phonon scattering.

A more elaborate calculation with higher-order phonon scatterings taken into account in the self-energy expression (10.2.3) endorses this. The convergence in summing the higher-order terms is so slow that one has to devise a method of extending the summation to infinity. One such calculation which considers all terms up to the sixth-order scatterings and a part of the higher-order (up to infinity) ones has been performed.¹⁵ The resulting lineshape is strongly asymmetric with a high-energy tail as is expected, with linewidth of the same form as (10.3.4) but with a constant factor 6.3 times as great.

The continuous transition of the spectral feature from the weak to the strong scattering limit, namely from the motionally-narrowed type to the local fluctuation type, as a function of D/B , has been studied by Sumi¹⁶ with the use of the coherent potential approximation (CPA, see Section 8.7). The fluctuation of local potential E_n is assumed to have a Gaussian distribution with dispersion D as is the case for localized excitation with linear interaction with a harmonic lattice described in Chapter 4, while the state density of the unperturbed exciton band is

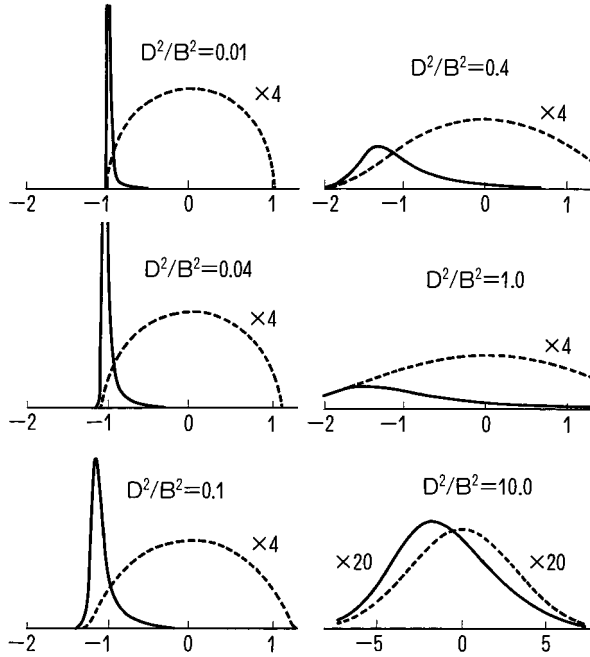


Fig. 10.2 Optical absorption spectra (solid lines) and density of states (dashed lines) of an exciton for various values of scattering strength $(D/B)^2$ which is proportional to temperature. Energy is measured from the band center in units of B . Due to Sumi.¹⁶

assumed to be of semi-spherical form with total bandwidth $2B$ as in eq. (8.9.1). The optically-allowed exciton state, $\mathbf{K} = 0$, is assumed to be at the bottom of this band. Figure 10.2 shows how the calculated spectra (solid line) and the density of states (dashed line) change as the fluctuation D increases against B (energy unit). If the abscissa is rescaled in units of D , one can see how the spectral width gets motionally narrowed as B increases against D (see the figure arranged in reversed order).

The above statements are based on the approximation of elastic scattering in which the phonon energies are neglected against the kinetic energy of the exciton. However, the finiteness of phonon energies removes the singular nature of the band bottom scattering mentioned above to a certain extent. The optically-created exciton at $\mathbf{K} = 0$ can be scattered to \mathbf{K}_{ac} or \mathbf{K}_{op} by absorbing an acoustic or optical phonon, with wave vector and energy conservation:

$$\hbar^2 K_{ac}^2 / 2m_{ex} = \hbar c_{\ell} K_{ac} \equiv \hbar \omega_{ac} = 2m_{ex} c_{\ell}^2 \equiv k_B T_{ac}, \quad K_{ac} \equiv 2m_{ex} c_{\ell} / \hbar, \quad (10.3.11)$$

$$\hbar^2 K_{op}^2 / 2m_{ex} = \hbar \omega_{\ell} \equiv k_B T_{op}, \quad K_{op} \equiv (2m_{ex} \omega_{\ell} / \hbar)^{1/2}. \quad (10.3.12)$$

The corresponding linewidths are given by

$$2\Gamma_{\text{ac}} = [\exp(T_{\text{ac}}/T) - 1]^{-1} [E_{\text{R}}^{(\text{ac})}/2\pi] (K_{\text{ac}}a_0)^3, \quad (10.3.13)$$

$$2\Gamma_{\text{op}} = [\exp(T_{\text{op}}/T) - 1]^{-1} [E_{\text{R}}^{(\text{op})}] (K_{\text{op}}a_0) \\ \times [\{1 + (\hbar\omega_{\ell}/4Ry_{\text{ex}})(m_{\text{h}}/m_{\text{e}})\}^{-2} \\ - \{1 + (\hbar\omega_{\ell}/4Ry_{\text{ex}})(m_{\text{e}}/m_{\text{h}})\}^{-2}]^2, \quad (10.3.14)$$

where $E_{\text{R}}^{(\text{ac})}$ and $E_{\text{R}}^{(\text{op})}$ defined in (9.5.6) (with $\Xi \equiv \Xi_{\text{e}} + \Xi_{\text{h}}$) have been used. The form factors have been set at unity for the acoustic mode since $K_{\text{ac}}a_{\text{ex}} \ll 1$.

The width (10.3.13) due to one-acoustic-phonon scattering becomes less than the width (10.3.7) at a temperature T given by

$$T > T_2 \equiv 0.01(K_{\text{ac}}a_0)B^2(E_{\text{R}}^{(\text{ac})})^{-1}k_{\text{B}}^{-1}. \quad (10.3.15)$$

Since the mechanism with the greater rate is to be considered the more appropriate, the multiphonon-scattering picture with elastic scattering is justified a posteriori at this high temperature. The criterion (10.3.15) is in fact consistent with the theoretical argument on the validity of elastic scattering.¹⁵ Itoh, Watanabe and Ueta measured the width of the excitonic molecule (instead of an exciton) in CuCl by giant two-photon absorption and analyzed its temperature dependence with the use of a one-phonon scattering rate at low temperature and a multiphonon-scattering rate at high temperatures.¹⁷ According to them $T_2 = 50$ K, $T_{\text{ac}} = 9.0$ K, $\Xi_{\text{mol}} = 2.6$ eV and $m_{\text{mol}} = 5.29$ times the free-electron mass.

Depending upon the matter, optical-phonon scattering may be the more important. Miyata¹⁸ found that the linewidths of the lowest excitons (which are well separated from their spin-orbit partners and expected to allow accurate linewidth analysis) in NaBr and NaI increase much more rapidly with temperature than would be extrapolated from low-temperature behavior through the coth-law based on the one-acoustic-phonon scattering model, and tentatively ascribed that to one-optical-phonon scattering.

10.4 Motional reduction of phonon sidebands and the zero-phonon line

Let us now study the phonon structures in the exciton spectra, confining ourselves to the absolute zero of temperature where the spectra can most clearly be resolved into a zero-phonon line and phonon sidebands of successive orders. Since the width function $2\bar{\Gamma}_0(E)$ for the lowest exciton (the suffix, $\lambda = 1$ s, will be omitted hereafter) at its renormalized energy $E = \bar{E}_0$ (where \bar{E}_K is the solution of $E = \bar{E}_K(E) \equiv E_K + \bar{\Delta}_K(E)$) vanishes as T tends to 0, one can write its normalized

lineshape as

$$\begin{aligned} \pi^{-1} \bar{I}_0(E) / [(E - \bar{E}_0)^2 + \bar{I}_0(E)^2] \sim [1 - d\bar{E}_0(\bar{E}_0)/dE]_I^{-1} [\delta(E - \bar{E}_0) \\ + \sum_{Ks} (H_I)_{0K}^{s-} (H_I)_{K0}^{s+} (\bar{E}_K + \hbar\omega_{s,-K} - \bar{E}_0)^{-2} \delta(E - \bar{E}_K - \hbar\omega_{s,-K})]_{II} \end{aligned} \quad (10.4.1)$$

where contributions from other exciton bands ($\lambda' \neq \lambda = 1s$) to \bar{I}_0 and \bar{A}_0 are neglected. The first and second terms in $[\cdots]_{II}$ represent the zero-phonon line and the one-phonon sideband, respectively, with their integrated intensities equal to the first and second terms in $[\cdots]_I$ thus assuring the normalization of the integrated spectrum of (10.4.1). (The equality of the second term of $[\cdots]_I$ and the integral of the second term of $[\cdots]_{II}$ can be confirmed with the use of the dispersion relation between the real and imaginary parts of the self-energy similar to (6.2.2, 6.2.3).)

Let us take the parabolic form of the renormalized exciton band

$$\bar{E}_K = \bar{E}_0 + \hbar^2 K^2 / 2\bar{m}_{ex} \quad (10.4.2)$$

with the renormalized effective mass \bar{m}_{ex} , and define

$$\hbar\Omega_{s,K} \equiv \hbar\omega_{s,K} + \hbar^2 K^2 / 2\bar{m}_{ex}. \quad (10.4.3)$$

Solving this for $\omega_{s,K}$, one puts

$$\omega_{s,K} = \omega_s^*(\Omega_{s,K}). \quad (10.4.4)$$

The functional relation between ω and Ω for acoustic and optical modes (s) can be read from Fig. 10.3. If the exciton were not mobile ($\bar{m}_{ex} \rightarrow \infty$), the second term of $[\cdots]_{II}$ would give the one-phonon sideband of a *localized* excitation, namely, \hbar^{-1} times the coupling function defined by (4.4.8):

$$s_s(\omega) = \sum_K (H_I)_{0K}^{s-} (H_I)_{K0}^{s+} (\hbar\omega_{s,K})^{-2} \delta(\omega - \omega_{s,K}). \quad (10.4.5)$$

The one-phonon sideband of the exciton line is given in terms of the coupling function

$$\ddot{s}_s(\Omega) = s_s(\omega_s^*(\Omega)) \cdot d\omega_s^*(\Omega)/d\Omega \cdot [\omega_s^*(\Omega)/\Omega]^2. \quad (10.4.6)$$

Equation (10.4.6) indicates that the coupling strength for the exciton is $[\omega_s^*(\Omega)/\Omega]^2 (< 1)$ times that for the localized excitation (as integrated). This is the “motional reduction” of the phonon sideband.⁹ As shown schematically in Fig. 10.4 for the typical case of a wide exciton band, the reduction is much more significant for the acoustic mode than for the optical mode. In this figure, the coupling function $s(\omega)$, shown by dashed lines, has been estimated for a localized electron at a lattice defect (or within an impurity atom) with s-like ground state of orbital radius r_s and

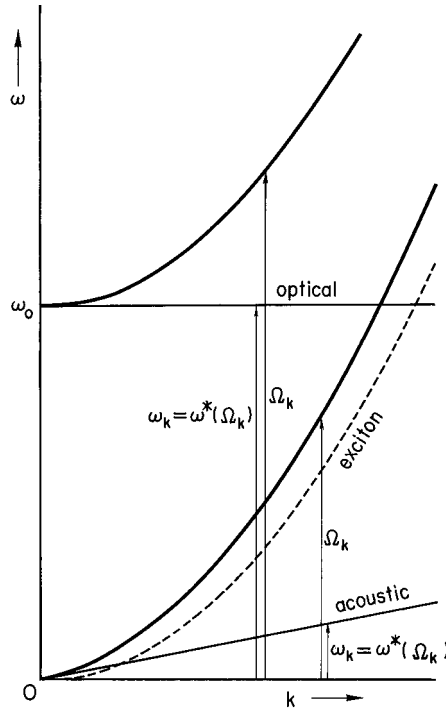


Fig. 10.3 Energies/ $\hbar = \Omega_k$ (thick solid lines) of the states consisting of an exciton (dashed line) and an acoustic phonon or an optical phonon (ω_K , solid lines) as functions of k .⁹ See text for a detailed explanation.

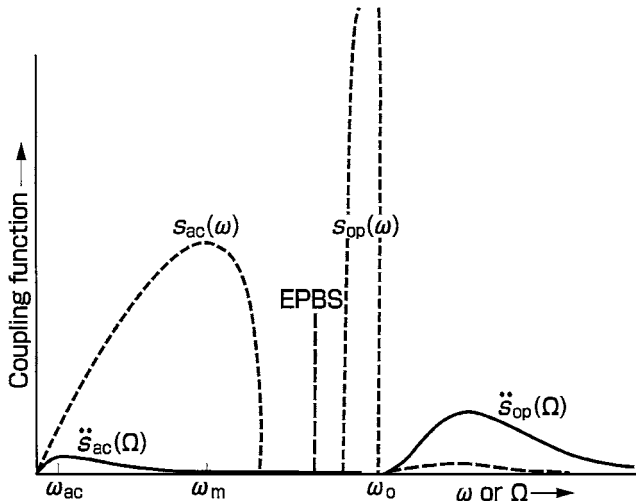


Fig. 10.4 Coupling function $\ddot{s}(\Omega)$ for an exciton (solid line) is compared schematically with the corresponding coupling function $s(\omega)$ for a localized excitation (short-dashed lines).⁹ For shallow excitons, exciton-phonon bound state (EPBS, long-dashed lines) may split off from the optical branch $\ddot{s}_{op}(\Omega)$.²⁰

p-like excited state with $r_p(\gg r_s)$ in a typical ionic crystal, with $S \equiv \int d\omega s(\omega)$ behaving as was shown in Fig. 4.15 as a function of $r_s(\gg a_0)$ (or $r_p(<a_0)$) (note that the electron radius closer to the lattice constant is more effective for the coupling strength S , see Section 4.8).

While $s_{ac}(\omega)$ rises linearly with ω , $\ddot{s}_{ac}(\Omega)$ reaches its maximum quite soon, at $\Omega \sim \overline{m}_{ex}c_l^2/2\hbar = \omega_{ac}/4$ (see (10.3.11)), and the motional reduction of the coupling strength given by \ddot{s}_{ac}/S_{ac} is of the order of $(\omega_{ac}/\omega_m)^2[\ln(\omega_m/\omega_{ac}) - 2]$ where ω_m is the maximum point of $s(\omega)$. Assuming (ω_{ac}/ω_m) to be of the order of 10^{-2} as is typically the case, one obtains a reduction factor $\sim 10^{-4}$. Since the energy separation between the zero-phonon line and the acoustic phonon sideband is also as small as 10^{-4} eV, they are unseparable (as is the case in most inorganic insulators and semiconductors) if the former is subject to inhomogeneous broadening of the same order due to crystal imperfections.

If the self-trapped state is more stable, the bottom of the free-exciton band which would be obtained by finite order renormalization is a metastable state with finite lifetime due to the quantum-mechanical tunneling to the former, and the true zero-phonon line is the optical transition to the self-trapped state (below the free state), with oscillator strength being reduced to $\exp(-S)$ times that of free state which is too small to be observed when $S > 10$ (the situation found in typical ionic crystals with self-trapping).

The strength of the two-acoustic-phonon sideband of an exciton which can be calculated with the higher-order terms included in (10.2.10) is not given by $\ddot{S}_{ac}^2/2!$ as in the localized excitation. In fact, it is much larger, larger even than \ddot{S}_{ac} , being of the order of $(\omega_{ac}/\omega_m)S_{ac}$ which can be comparable to the zero-phonon line.

In this way, the multiphonon structures in exciton spectra are completely different from those of localized excitation as far as the acoustic branch is concerned, except for those molecular crystals with exciton transfer energy as small as the average acoustic-phonon energy. Except at very low temperatures, the acoustic-phonon structures do not show up in exciton spectra and, in most cases, the overall lineshapes have been analysed without decomposing them into multiphonon structures.

The difference between the multiphonon structures of localized excitation and those of an exciton is less drastic for the optical branch. Assuming a small electron-hole effective mass ratio, one can estimate the motional reduction factor: $\ddot{S}_{op}/S_{op} \sim r_h \cdot k_p \sim O(10^{-1})$ where r_h is the orbital radius of the hole in the exciton and k_p is the reciprocal polaron radius defined by (9.1.16). In fact, an optical-phonon sideband has been observed at an energy somewhat above that of the 1s exciton plus one optical phonon, with strength \ddot{S}_{op} an order of magnitude smaller than S_{op} estimated for the corresponding localized excitation, as seen in Fig. 10.5 which is the spectra of NaBr due to Miyata.¹⁸ The 2s and higher excitons with larger radii can be scattered only by longer-wavelength phonons, which means that the effective recoil kinetic

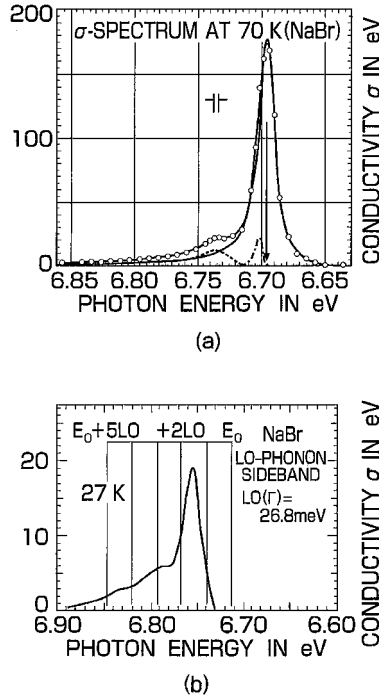


Fig. 10.5 (a) Conductivity (σ) spectrum of NaBr in the energy region of the Γ ($3/2$, $1/2$)-exciton. (b) LO-phonon sidebands on expanded scale. Due to Miyata.¹⁸

energy of these excitons is smaller. Namely, the larger-radii excitons behave more like the localized excitation as far as the phonon sidebands are concerned. This is consistent with the observation of several optical-phonon sidebands of 2s excitons in the reflection spectrum of KI with intervals almost equal to the optical-phonon energy due to Baldini *et al.*,¹⁹ as seen in the inset of Fig. 10.6. A weak hump on the high-energy side of the 1s-exciton peak, which was tentatively ascribed to the localized mode optical phonons because of their greater splitting energy, is now to be ascribed to the optical-phonon sideband of the *recoiled* exciton as mentioned above. These features of small- and large-radii excitons are consistent also with the theoretical studies of the exciton polaron due to Matsuura and Büttner⁵ mentioned in Section 10.1.

Depending upon the matter, there can arise a bound (or quasi-bound) state of an exciton and an optical phonon (EP(Q)BS) due to the internal states of the former resonating with the latter, the mechanism proposed^{20,21} to explain the absorption peak somewhat below the energy of the 1s exciton plus one optical phonon observed in some ionic crystals, such as ZnO,²² MgO and BeO,²³ TlCl and TlBr.^{24,25} This is an example of a final-state interaction: the interaction between two particles simultaneously created by optical excitation results in energy structures different

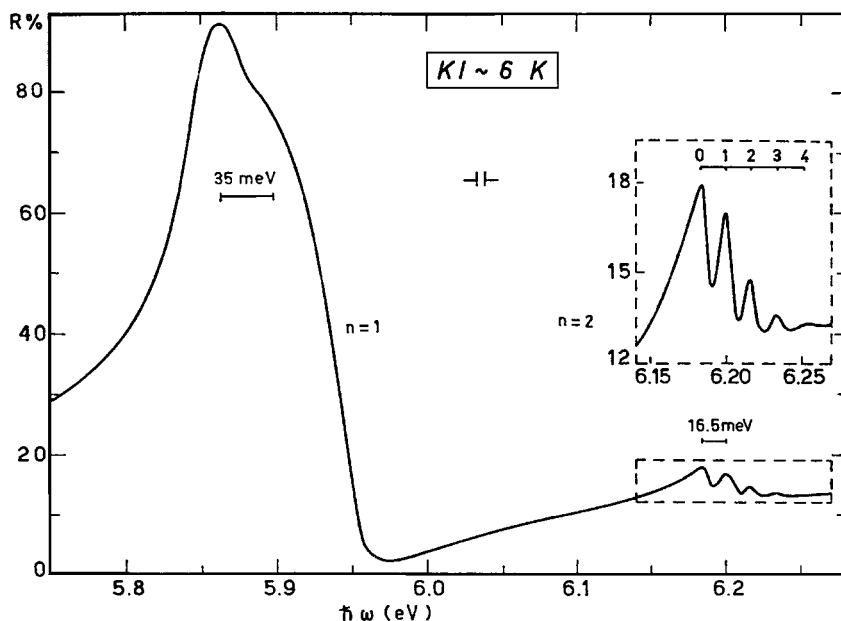


Fig. 10.6 Reflectivity spectrum of 1s- and 2s-exciton peaks of KI. The 2s exciton is also indicated on expanded scale in the inset. Due to Baldini *et al.*¹⁹

from the sum of the energies of the two particles, with possible formation of bound states. An exciton itself is the simplest example of a final-state interaction between the optically-created electron and hole.

Coming back to the general formula (10.2.12), *both* the main part of the exciton line and its phonon sidebands are incorporated into a *single* Lorentzian expression, the former through the value of the width function $\bar{\Gamma}_{\lambda 0}(E)$ at the renormalized energy $E = \bar{E}_{\lambda 0}$ as described in Section 10.4, and the latter through the energy dependence of $\bar{\Gamma}_{\lambda 0}(E)$ as has been described above. This feature is clearly seen in Figs. 10.5 and 10.6.

10.5 Phonon sidebands of vibronic excitons in molecular crystals

We have so far considered crystals in which the exciton bandwidth is much greater than the phonon energy. The situation is different in those molecular crystals in which the intermolecular transfer energy t_c of intramolecular excitation is smaller partly due to the smaller overlap of molecular wave functions. Moreover, the intramolecular excitation consists of a series of *vibronic* states due to the intramolecular vibrations. The normalized intensity of the n th vibronic line is given by the Poisson distribution: $I_n = \exp(-S)S^n/n!$ (the squared overlap integral of the n th vibrational state in the electronic excited state with the lowest vibrational state

of the electronic ground state) in the case of one harmonic oscillator with linear interaction with the electron through the coupling strength S as was described in Chapter 4. Since the interatomic bonds within a molecule are much stronger than the intermolecular bonds, a quantum of an intramolecular vibration is much greater than the typical phonon energy of the lattice vibrations. Hence, it is more realistic to consider *each vibronic* state as an excited state which forms its own Frenkel exciton band, often called a “vibronic exciton”. The transfer energy of a vibronic exciton is given by $t = t_e \exp(-S) S^n / n!$ since the squared transition dipole moment of a molecule is reduced by the factor I_n . The transfer energy t_e which itself is small is thus further reduced by more than an order of magnitude for the vibronic-exciton band if S is greater than 3. If t thus reduced is smaller than the phonon energies, the phonon sidebands are more similar to those of the localized excitation, and the recoil kinetic energy of the exciton plays only a subsidiary role.

As a model material with such vibronic excitons, we will take NaNO_2 , which has been studied most systematically by Kato and his collaborators.^{26–31,33} It is an orthorhombic body-centered crystal, ferroelectric with dipoles of bended NO_2^- molecules aligned along the b -axis below the order–disorder transition temperature $T_c = 163^\circ\text{C}$. The uv absorption, attributed to the electronic transition $^1A_1 \rightarrow ^1B_1$ with transition dipole moment perpendicular to the molecular plane of the NO_2^- molecule and along the a -axis of the crystal, consists, at very low temperatures, of a number of vibronic lines: $\bar{\nu}_{n_1, n_2}^1 = \bar{\nu}_{00} + n_1 \bar{\nu}_1^1 + n_2 \bar{\nu}_2^1$ with non-negative integers n_1 and n_2 where $\bar{\nu}_{00} = 25975 \text{ cm}^{-1}$ (zero-phonon line), $\bar{\nu}_1^1 = 1020 \text{ cm}^{-1}$ and $\bar{\nu}_2^1 = 630 \text{ cm}^{-1}$ according to Kamada, Yoshikawa and Kato,²⁶ as shown in Fig. 10.7(a). (Note that $\bar{\nu}$ represents the wave number, per unit length, of light in a vacuum, ν/c_0 , the conventional unit in spectroscopy.) Correspondingly, the luminescence spectrum consists of vibronic lines, $\bar{\nu}_{n_1, n_2} = \bar{\nu}_{00} - n_1 \bar{\nu}_1 - n_2 \bar{\nu}_2$ where $\bar{\nu}_1 = 1328 \text{ cm}^{-1}$ and $\bar{\nu}_2 = 826 \text{ cm}^{-1}$ according to Kamada, Yoshimura and Kato.²⁷ The wave numbers $\bar{\nu}_1$ ($\bar{\nu}_1^1$) and $\bar{\nu}_2$ ($\bar{\nu}_2^1$) correspond respectively to the stretching and bending modes in the ground (excited) state of the NO_2^- molecule. Several vibronic lines of the latter mode are observed in the absorption and luminescence spectra, with coupling strength $S_2^1 \sim 3$ and $S_2 \sim 4$ as estimated from the intensity ratio of the successive vibronic lines.

The last fact is consistent with the configuration coordinate model.²⁷ Modifying Fig. 4.4 so as to consider the difference in the vibrational frequency between the ground and excited states, one can construct the configuration coordinate model for the bending mode q_2 with relative displacement Δ in the equilibrium position as shown in Fig. 10.8. According to Section 4.5, the relaxation energy defined as the difference of the adiabatic energy between the Franck–Condon state (immediately after the optical transition) and the relaxed state can be written, for the ground and

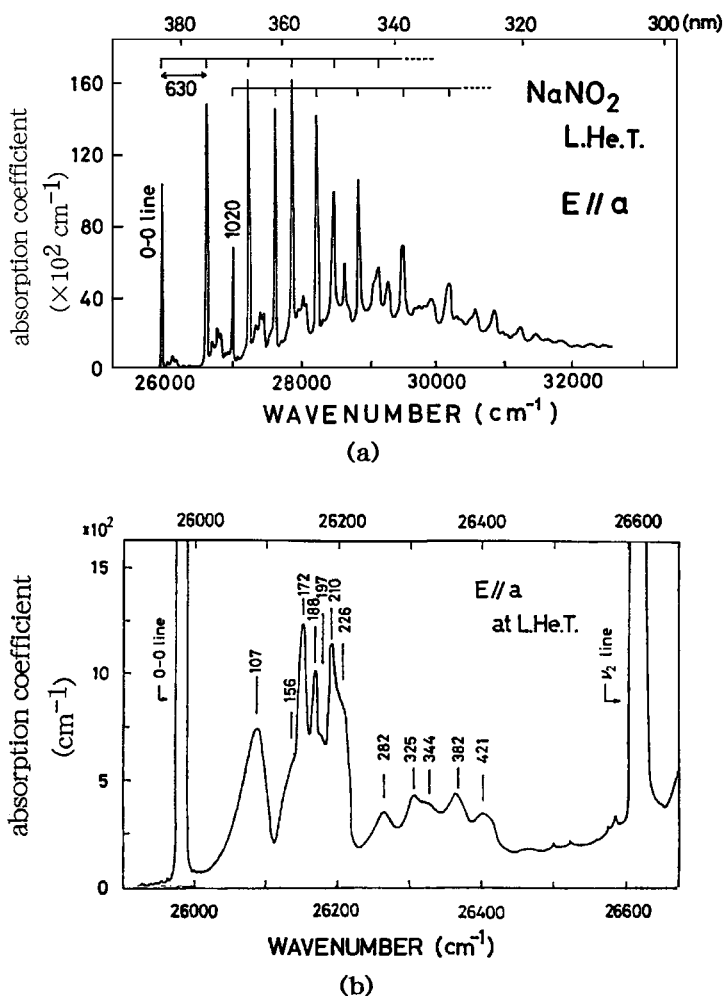


Fig. 10.7 (a) The uv absorption spectra of NaNO_2 polarized along the crystallographic a -axis. (b) One-phonon sidebands near the zero-phonon line on expanded scale. Due to Kamada *et al.*²⁶

excited states, as

$$E_R^{(g)} = (2\pi\nu)^2 \Delta^2 / 2 \equiv Sh\nu, \quad E_R^{(e)} = (2\pi\nu')^2 \Delta^2 / 2 \equiv S'h\nu'. \quad (10.5.1)$$

Exactly speaking, the intensity distribution among successive vibronic lines is not a Poisson distribution when $\nu \neq \nu'$. However, each vibrational state has its greatest amplitude at its classical-mechanical turning point (where the particle spends the longest time), and that vibrational state which has its turning point at the Franck–Condon state has the greatest overlap with the lowest vibrational state

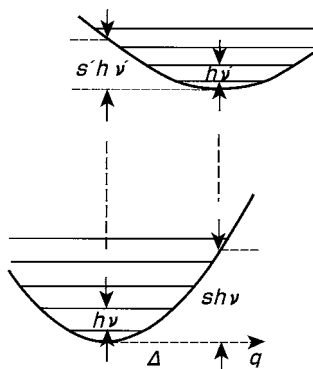


Fig. 10.8 Configuration coordinate model of a molecule, with displaced (by Δ) equilibrium positions and with *different* vibrational frequencies (ν , ν') of the ground and excited electronic states.

before the optical transition. Hence, the fact that the vibronic line has the greatest intensity at $n(n') = S(S') - 1/2$, namely that $S(S')$ is given by the peak vibronic line n plus $1/2$, holds even if $\nu \neq \nu'$. The observed ratios $S'/S = 3/4$ and $\nu'/\nu = 0.76$ are in satisfactory agreement with the relation $S'/S = \nu'/\nu$ obtained from (10.5.1).

Each vibronic line of the absorption (luminescence) spectra is accompanied on its high- (low-) energy side by a phonon sideband which is contributed by three acoustic, three optical and three (molecular) librational modes, as is seen more clearly in the expanded Fig. 10.7(b). The structures in the phonon sideband of the 00-line in the absorption spectra have been analysed on the basis of the energy-momentum conservation of the exciton-phonon system: $E = \varepsilon(\mathbf{q}) + h\nu(-\mathbf{q})$ where E is the photon energy, and with the help of the dispersion curves of phonons due to the neutron diffraction studies by Dolling, Sakurai and Cowley.²⁸ The dispersion of the 00-vibronic excitons, which turned out to be small compared with that of phonons, was further studied through the secondary emission under energy selective excitation by Sakai *et al.*²⁹ and by comparison of the phonon sidebands in the absorption and luminescence spectra by Ashida *et al.*³⁰ The results indicate that the width of the 00-exciton band is 3–5 cm^{-1} , more than an order of magnitude smaller than the typical phonon energies of a few hundreds cm^{-1} . These studies revealed also that at low temperatures the exciton cannot be thermalized, before its radiative annihilation, in this narrow band where the one-phonon absorption and emission processes satisfying energy-momentum conservation cannot take place. Time-resolved study of the lineshape of $\bar{\nu}_{0,2}$ luminescence spectra by A. Kato *et al.*³¹ revealed that the thermalization time is proportional to T^{-7} , indicating the dominance of two-acoustic-phonon scattering studied by Holstein, Lyo and Orbach.³² These situations are never met in wide-band excitons. A zero-phonon

line separated from the acoustic phonon sideband, which has never been observed in wide-band excitons due to the motional effect, is in fact observed in this vibronic exciton due to its extremely small dispersion (absence of the motional effect).

If the self-trapped state is more stable than the free state in materials with a wide exciton band, as was mentioned in Sections 9.5 and 9.6 with many examples, the genuine zero-phonon line is the transition to the former state, which has never been observed. The reason is obviously the reduction factor $\exp(-S)$, with $S \sim$ several tens for alkali halides with exciton radius of the order of the lattice constant as can be read from Fig. 4.15. For the same reason, the enormous effective mass $m^* = m_{\text{ex}} \exp(S)$ of the self-trapped exciton allows one to imagine a practically immobile localized entity.

In comparison with the self-trapped state in wide-exciton-band materials, one may well ask about the contribution S_{ph} of the phonon sideband to the reduction factor of the bandwidth of the 00-vibronic exciton in NaNO_2 . The total coupling strength S consists of the contributions S_1 and S_2 from intramolecular vibrations (the third intramolecular vibration, the antisymmetric stretching mode, has no linear coupling so that $S_3 = 0$) and S_{ph} . The oscillator strengths analysed from the reflection spectra³³ leads to $S_2 \sim 3$ and $S_1 \sim 0.4$. The one-phonon sideband is continuous and cannot be well separated from its multiple sidebands consisting of the convolutions of the former (see (4.5.16)). Their total intensity, $\sum_{n=1}^{\infty} (S_{\text{ph}})^n / n! = \exp(S_{\text{ph}}) - 1$, is roughly estimated to be $1.5 \sim 4$ times S_2 from the comparison of the area of the continuous phonon sidebands with that of the vibronic lines $n_2 = 1$ in Fig. 10.7, and hence $S_{\text{ph}} = 1.7 \sim 2.6$. Thus one estimates $S = 5 \sim 6$ and the total reduction factor of the 00-line intensity and hence of the 00-exciton band width as $t/t_e = \exp(-S) = 0.007 \sim 0.002$. The moderate value of S_{ph} in NaNO_2 compared to the large value of S (several tens) in alkali halides is due to the orbital radii of the ground- and excited-state electrons within a NO_2 molecule which are small compared to the lattice constant (see the left pieces of the curves in Fig. 4.15), while the not-too-large values of S_1 and S_2 are due to the high frequency ν of the intramolecular vibrations ($S = E_R / h\nu$). Thus, S for NaNO_2 is not so large as to make the zero-phonon line invisible, and the 00-vibronic exciton band is to be called *narrow band* rather than self-trapped state. The difference in S is greatly magnified in $\exp(S)$, resulting sometimes in qualitatively different situations.

10.6 Correlation of absorption and emission spectra of excitons in the phonon field

10.6.1 The Urbach rule

Since the discovery of this empirical rule by Urbach³⁴ in silver halides, it has been established for a variety of insulators – organic as well as inorganic³⁵ – that the

lower-energy side of the first exciton absorption peak decays exponentially as energy decreases, with a decay constant proportional to the reciprocal temperature at high enough temperatures:

$$F_a(E) = A \exp[-\sigma(E'_0 - E)/k_B T], \quad (10.6.1)$$

as shown in Fig. 10.9 for KI crystal due to Haupt.³⁶ The convergence point, E'_0 , of the extrapolated semilogarithmic plots of the absorption spectra at various temperatures is usually situated near the absorption peak at the lowest temperature. The dimensionless “steepness coefficient”, σ , is a material constant of the order of unity. The exponential form of (10.6.1) holds over a few to several decades in $F_a(E)$. The absorption spectra at low temperatures are more varied: in some materials the exponential energy dependence holds approximately with $k_B T$ being replaced by $(\hbar\omega/2) \coth(\hbar\omega/2k_B T)$ with appropriate angular frequency ω , while in other materials the spectra show some deviation from the exponential form.

The universality and simplicity of this empirical rule at high temperatures – now called the Urbach rule – have evoked a number of theoretical and experimental studies to explain it on a microscopic basis. It is difficult to review these studies, which are so varied and diverse, within a reasonable space. We will trace only a

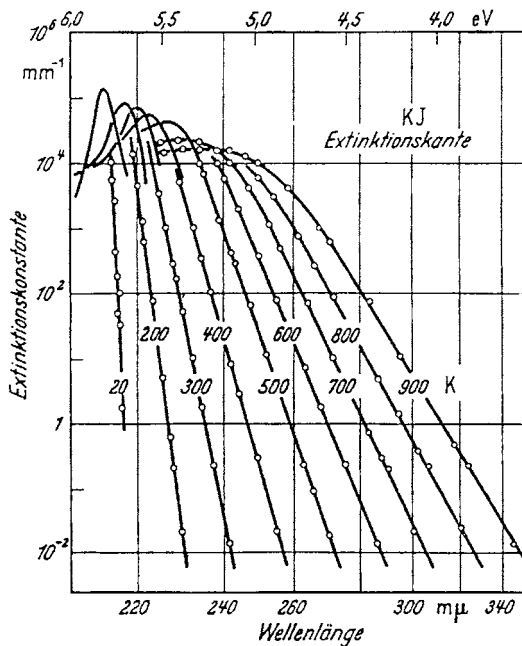


Fig. 10.9 Semilogarithmic plot of the attenuation coefficients (solid lines) of KI crystal versus photon energy at various temperatures. Due to Haupt.³⁶

few of them from our viewpoint, referring to Refs. [35] and [37] for more complete references.

First, in view of its temperature dependence, the exponential tail is undoubtedly an effect of lattice vibrations on the electronic state, as was discussed by Dexter.³⁸ Second, it is almost obvious, in view of the absence of phonon structures in this tail part (except possibly at low temperatures where this rule no longer holds), that the adiabatic and classical treatment of lattice vibrations is enough for its explanation.

A tentative microscopic model to derive the empirical rule (10.6.1) analytically was proposed, in which an exciton was supposed to couple linearly with an even mode, and quadratically with an odd mode, of the lattice vibrations,³⁹ although such an adiabatic picture was admitted to be applicable to a localized exciton but not to a free exciton due to its recoil scattering by phonons. This two-mode model was revived by Mahr⁴⁰ to explain his data on the absorption band due to an impurity-bound exciton which is Gaussian in the central part (ascribed to linear coupling) and of the Urbach form in the lower-energy tail (ascribed to quadratic coupling). His argument is legitimate since a *bound* exciton is *not recoiled*.

In view of the apparent similarity between the Urbach rules for intrinsic and extrinsic excitons, another attempt was made to apply the two-mode model to an intrinsic exciton, this time not directly to its free states but to the “momentarily localized states” (which vary from time to time, adiabatically following the lattice vibrations).⁴¹ This notion was introduced, not only for the ad-hoc reason of being recoilless, but also due to the realization¹⁵ that the successive perturbation theory starting from free-exciton state is an extremely roundabout approach to the explanation of the exponential tail. The existence of localized states below the mobile states in a randomly perturbed lattice was also widely realized due to the famous paper by Anderson.⁴² However, it seemed extremely difficult to give a quantitative study of the configurational statistics for momentarily localized states. Although a correlation between the width of the tail – the reciprocal of the steepness coefficient σ – and the exciton–phonon coupling constant g (introduced in Section 9.5 and to be redefined later) was noted in these qualitative derivations of the Urbach rule, and hence a correlation between the absorption (unrelaxed exciton) and emission (relaxed exciton) spectra was recognized, it was desired to find a theoretical framework to treat quantitatively the momentarily localized states induced by the spacially fluctuating field of the lattice vibrations.

Substantial progress was made by the application of the average t-matrix approximation (ATA)⁴³ and the coherent potential approximation (CPA)⁴⁴ which had been developed to study random lattice problems as already mentioned in Section 8.7. In these approximations, the localized states are directly and properly incorporated in the self-energy. As a result, the low-energy tail of the Gaussian distribution of site energies manifests itself as a pseudo-exponential tail of the absorption spectrum

resembling (10.6.1) in its energy and temperature dependence. From a quantitative point of view, however, there remained a non-trivial discrepancy. In particular, the fact that the lineshape is convex just below the peak seems to originate from the well-known limitation of ATA and CPA that coherent scattering by more than two sites is not taken into account. A correction to this underestimation seemed desirable to improve the calculated lineshape towards (10.6.1), because the shallowly localized states due to coherently negative potentials would contribute giant oscillator strength to this spectral region.

A different mechanism was proposed by Dow and Redfield⁴⁵ who ascribed the exponential tail to the Franz–Keldysh type broadening due to exciton ionization under the fluctuating electric field caused by the lattice vibrations. What is concerned here is the electron–hole relative motion, instead of the translational motion considered thus far. An experimental study of the effect of a static electric field on the exponential tail⁴⁶ seemed to favor this theory. A difficulty with the theory is the unrealistic assumption of a uniform electric field in calculating the ionization rate of a large-radius exciton although the statistical distribution of the electric fields is considered carefully. Moreover the Urbach rule has been observed in non-polar crystals as well, in which the lattice vibrations are not accompanied by strong electric fields.

10.6.2 A model Hamiltonian for an exciton in a deformable lattice – fluctuation and relaxation

Coming back to the translational motion, the only tractable method to overcome the remaining difficulty inherent to the one-site approximations seemed to be to resort to a direct numerical calculation, which was performed by Schreiber and Toyozawa.³⁷ Although the problem of exciton–phonon interaction has been treated in various ways in the preceding sections, it is useful, for our purpose of correlating the absorption (fluctuation) and emission (relaxed state) spectra, to simplify the model Hamiltonian H into the following form:^{14,47}

$$H = H_e + H_I + U_L + K_L, \quad (10.6.2)$$

where

$$H_e = \sum_n |n\rangle E_a \langle n| + \sum \sum_{n \neq m} |n\rangle t_{nm} \langle m|, \quad (10.6.3)$$

$$H_I = - \sum_n |n\rangle c Q_n \langle n|, \quad (10.6.4)$$

$$U_L = \sum_n Q_n^2 / 2, \quad K_L = \sum_n (\omega^2)_{nn'} P_n P_{n'} / 2. \quad (10.6.5)$$

In (10.6.3) we consider an exciton in the tight-binding picture, with intrasite excitation energy E_a and intersite transfer energy t_{nm} . (The electron–hole relative motion

is ignored since we are here to ascribe the exponential tail to the translational motion of an exciton in the fluctuating field.) The eigenstates of H_e consist of the exciton band with eigenenergy given by

$$E_{\mathbf{K}} = E_a + \sum_m t_{0m} \exp(i\mathbf{K} \cdot \mathbf{R}_m). \quad (10.6.6)$$

Denote by B the energy difference between the band center given by $N^{-1} \sum_{\mathbf{K}} E_{\mathbf{K}}$ and the band bottom which is at $\mathbf{K} = \mathbf{0}$ under the assumption $t_{0m} > 0$ (for all m),

$$B \equiv E_a - E_0 = - \sum_m t_{0m}. \quad (10.6.7)$$

But for the exciton-phonon interaction H_I , the optical transition from the ground state is allowed only to the state $|\mathbf{K} = \mathbf{0}\rangle$ of the exciton band due to the \mathbf{K} -selection rule. The normalized absorption spectrum is then given by

$$F_a(E) = \delta(E_0 - E), \quad (10.6.8)$$

as shown schematically in Fig. 10.10(a) and (a') for small and large $|t|$, respectively. For definiteness, we will assume a simple lattice with transfer t only to the $2d$ nearest neighbors where $d(=1, 2, 3)$ is the dimensionality of the lattice. Then the state density is symmetric around the center, with total band width $2B$.

The fluctuation of site energy, given by (10.6.4), originates from lattice vibrations with Hamiltonian $H_L = K_L + U_L$ given by (10.6.5). The local vibrational coordinates $\{Q_n\}$ are so chosen that U_L , instead of K_L , is diagonal and the frequency ω appears in the kinetic energy K_L which is non-diagonal. (The reasons for this unconventional choice will be mentioned later.)

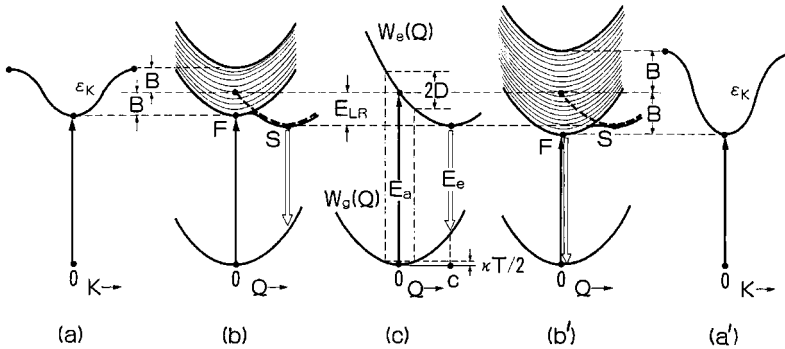


Fig. 10.10 Configuration coordinate models for localized excitation (c) and the exciton [(b) or (b')], and the dispersion of the exciton band in the rigid lattice [(a) or (a')]. The strong and weak coupling cases ($B \lesseqgtr E_R$) are represented by (a), (b) and (a'), (b'), respectively.^{41,47}

In the absence of transfer ($t_{nm} = 0$), the localized excitation (say, $|n\rangle$) gives an eigenstate, with adiabatic potential

$$\begin{aligned} W_{en}(Q) &= \langle n | H_e + H_I | n \rangle + U_L(Q) = E_a - cQ_n + \sum_{n'} Q_{n'}^2/2 \\ &= E_a - E_R + (Q_n - c)^2/2 + \sum_{n'(\neq n)} Q_{n'}^2/2, \end{aligned} \quad (10.6.9)$$

$$\text{where } E_R \equiv c^2/2. \quad (10.6.10)$$

The adiabatic potential of the ground state is simply given by

$$W_g(Q) = U_L(Q). \quad (10.6.11)$$

After Franck–Condon excitation of the n th site (vertical transition in Fig. 10.10(c)) at $Q = 0$, the lattice will relax to a new equilibrium at $Q_n = c$, $Q_{n'} = 0$ ($n' \neq n$), thereby releasing the lattice relaxation energy E_R to other sites through K_L of (10.6.5) even under the assumption $t = 0$. Let us take into account the effect of lattice vibrations on the absorption spectra by assuming the Boltzmann distribution: $\exp[-W_g(Q)/k_B T]$ in the initial state. Applying the Franck–Condon principle with a Q -independent transition dipole moment, we can calculate the normalized absorption spectra as

$$\begin{aligned} F_a(E) &= \int dQ_1 \cdots \int dQ_N \delta[W_{en}(Q) - W_g(Q) - E] \exp[-W_g(Q)/k_B T] \\ &\quad \div \int dQ_1 \cdots \int dQ_N \exp[-W_g(Q)/k_B T] \end{aligned} \quad (10.6.12)$$

$$= (2\pi D^2)^{1/2} \exp[-(E - E_a)^2/2D^2], \quad (10.6.12a)$$

$$\text{where } D^2 \equiv \langle V_n^2 \rangle = 2E_R k_B T. \quad (10.6.13)$$

The amplitude of the local energy fluctuation $V_n = -cQ_n$ is reflected in the Gaussian width D . The luminescence spectrum can be calculated in the same way if thermal equilibrium within the local excited state $|n\rangle$ is reached before the emission process; the spectral shape is the same as (10.6.12a) with E_a replaced by $E_e = E_a - 2E_R$. Here Q_n is a sort of the *interaction mode* which was defined for the relaxation of localized excitation in Section 4.4.

It should be noted in passing that H_I , U_L and hence the relaxation energy E_R of a localized exciton are quantities of *electronic* origin, being independent from the *atomic masses* (as they should) which are relevant only to the kinetic energy K_L through ω . What we called an unconventional choice of Q and P in the above is physically a natural choice. The second reason for this choice is that K_L with ω is relevant to the fine structures of spectra due to phonons but not to the overall lineshape concerned here for which the classical description of lattice vibrations with the use of the adiabatic *potential* is enough.

In the general case with $c \neq 0$ and $t \neq 0$, the configuration coordinate model corresponding to Fig. 10.10(c) is obtained by incorporating the continuum of the exciton band of Fig. 10.10(a) or (a') into (c), as shown schematically in Fig. 10.10(b) or (b').^{41,47} As we switch on the local distortion Q_n and hence the local potential $-cQ_n$ at the cost of lattice energy $Q_n^2/2$, the energy levels of extended exciton states are subject to an infinitesimal change of $O(N^{-1})$, being parallel to the ground-state parabola, while a bound state splits off out of this continuum, possibly forming another minimum S. The state S is nothing but the self-trapped state – the exciton is stabilized by being trapped by the lattice distortion it has induced. With increase of Q_n , this discrete level asymptotically approaches the broken line which is a reproduction in Figs. 10.10(b) and (b') of the $W_e(Q)$ line (completely localized excitation) in Fig. 10.10(c) with the same value of c . Neglecting the small energy difference between the solid and broken lines at $S(\sim |t|^2/2E_R = B^2/2zE_R$, minus the second-order perturbation energy of the latter, if the transfer energy, t , to the z nearest neighbors is considered), one has a lower- or higher-energy than F (the lowest free state without lattice distortion) according as $B < \text{or} > E_R$, corresponding to the case (b) or (b'). The existence of a potential barrier between the free state F at the bottom of the exciton band and the self-trapped state S (due to the short-range exciton–phonon interaction in the three-dimensional lattice as stated in Sections 9.4 and 9.5) is schematically incorporated in (b') and (b). Therefore, the stable exciton state in the phonon field changes abruptly from F- to S-state as the exciton–phonon coupling constant, defined by

$$g \equiv E_R/B, \quad (10.6.14)$$

exceeds $g_c \equiv 1 - (2z)^{-1} \sim 1$ (note that $(2z)^{-1} = 0.083$ for a simple cubic lattice). The exciton usually stabilizes itself in the lattice (in $\sim 10^{-12}$ s) well before it emits a photon (in $\sim 10^{-9}$ s). Hence, the emission band appears at $E_a - B = E_0$ (F-type: sharp line, resonant to the absorption band; Fig. 10.10(b')) or at $E_a - 2E_R$ (S-type: broad, Stokes-shifted band; Fig. 10.10(b) according as $g < \text{or} > g_c$, as also shown in the lower half of Fig. 10.11.¹⁴ The feature of the emission band changes *abruptly* when g/g_c exceeds unity.

The upper half of Fig. 10.11 describes the feature of the absorption spectra which is governed by the competition of D and B and varies gradually from weak to strong scattering as D/B increases. As can be seen by comparison of the upper and lower parts of the figure, the region $D/B \sim 1$ of *gradual* transition between weak and strong *scattering* is different from the boundary $E_R/B = g_c(\sim 1)$ of *abrupt* transition between weak and strong *coupling*, in such a way that there is a region with *strong coupling but weak scattering* provided $k_B T \ll B$. The configuration

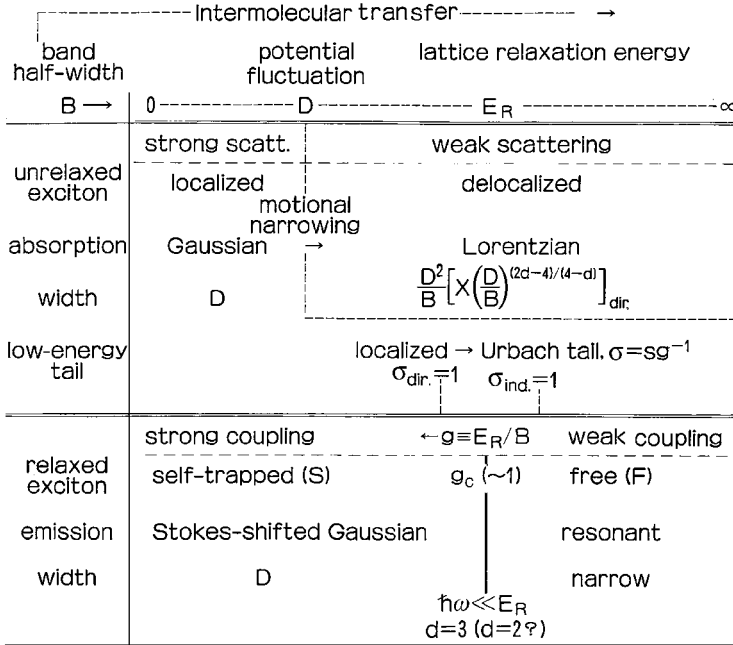


Fig. 10.11 Perspective on localization versus delocalization of an exciton in the phonon field as reflected in the optical spectra. Note that B increases towards the right with D and E_R being fixed.¹⁴

coordinates $\{Q_n\}$ relevant to the absorption spectra are confined to thermal vibrations with amplitude much smaller than the displacement c in the relaxed excited state. The normalized lineshape of the absorption spectra is given by

$$F_a(E) = \sum_i \overline{\langle \mathbf{0} | i(Q) \rangle \delta[E - E_i(Q)] \langle i(Q) | \mathbf{0} \rangle} \quad (10.6.15)$$

where $|\mathbf{0}\rangle$ means the free-exciton state with $\mathbf{K} = \mathbf{0}$, $|i(Q)\rangle$ and $E_i(Q)$ are eigenstates and eigenenergies of $H_e + H_l(Q)$ for a configuration $Q \equiv (Q_1, Q_2, \dots, Q_N)$ and $\overline{\dots}$ indicates to take the average over the Boltzmann distribution of Q as was done in (10.6.12). Thus we have to solve a typical eigenvalue problem of a random lattice where the electron (exciton) is subject to a spacially random potential $V_n = -cQ_n$ obeying the Gaussian distribution $\exp(-V_n^2/2D^2)$ without mutual correlation: $\overline{V_n V_m} = 0$. Instead of the integration (10.6.12), the summation on the Monte Carlo distribution of (Q_1, Q_2, \dots, Q_N) was taken for the calculation of (10.6.15) in Ref. [37]. A linear chain ($d = 1$) of 30 sites, a square lattice ($d = 2$) of 13×12 sites and a simple cubic lattice ($d = 3$) of $10 \times 9 \times 8$ sites were taken.

10.6.3 Calculated lineshapes – the central part

Typical lineshapes calculated³⁷ for 1-, 2- and 3-dimensional lattices for the case of $g \equiv E_R/B = 1$ and at typical temperatures well within the weak scattering regime are shown in Fig. 10.12 by solid lines. For comparison, the Lorentzian curve fitted to the high-energy side of the peak is also shown by a broken line. The abscissa denotes the photon energy E measured from the band center E_a as origin with energy unit $2B = 1.0$ eV, the total bandwidth (so that the unrenormalized band bottom is at -0.5). The most remarkable feature is the strong asymmetry (due not to the \bar{A} -term but to the E dependence of the width function $\bar{\Gamma}(E)$ in (10.2.12)) with steeper drop on the lower-energy side, indicating far fewer eigenstates E_i below the renormalized band bottom (\sim peak position). On the opposite side of the peak, the sign of the discrepancy of the calculated lineshape from the Lorentzian with increasing energy depends on dimensionality. This is to be ascribed to the energy dependence of $\bar{\Gamma}(E)$ (proportional to the renormalized density of states) which decreases, stays constant or increases as E increases above the renormalized bottom of the exciton band according as $d = 1, 2$ or 3 , as is well known. In fact, if one calculates the renormalized density of states

$$\rho(E) = N^{-1} \sum_i \overline{\delta[E - E_i(Q)]} \quad (10.6.16)$$

and the *average oscillator strength per state* (AOSPS) defined by

$$f(E) \equiv F_a(E)/\rho(E) = N[\overline{|\langle \mathbf{0} | i(Q) \rangle|^2}]_{E_i(Q)=E}, \quad (10.6.17)$$

one finds that $f(E)$ fits almost exactly to a Lorentzian curve in the weak scattering regimes in accordance with our perturbation theory interpretation of the linewidth as (scattering-limited) lifetime broadening. This behavior of $F_a(E)$ and $f(E)$ will

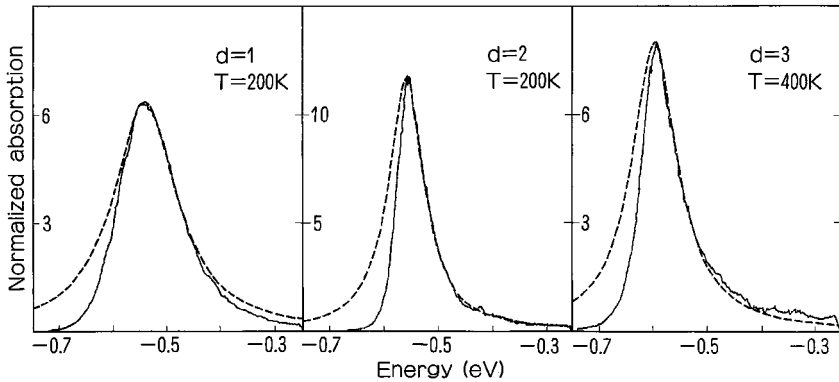


Fig. 10.12 Calculated exciton absorption spectra (solid line) for three dimensionalities, compared with Lorentzian (broken line) best-fitted to the high-energy side of the peaks.³⁷

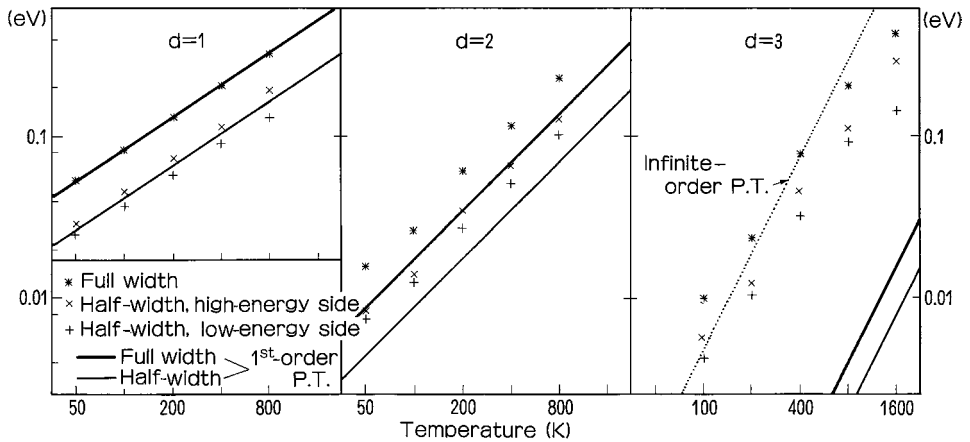


Fig. 10.13 Calculated linewidths (symbols) are plotted against temperature and compared with perturbation theory (lines).³⁷

be analyzed in more detail in (Subsection 10.6.4) with particular attention to the different origins of the high- and low-energy tails.

The widths of the calculated lineshapes are plotted against temperature in Fig. 10.13, and are in good agreement with the perturbation theory calculations described before, especially in that the temperature dependence is given by $T^{2/(4-d)}$. The deviations from this in the $d = 3$ case at low and high temperatures are due, respectively, to the number of lattice sites considered – too few to calculate this small width – and to the onset of the intermediate scattering regime. It should also be noted that the inclusion of higher-order scattering terms in the perturbation expansion is essential to obtain the right order of magnitude of the linewidth.

Quantitative analysis of the observed lineshapes of exciton absorption spectra has been performed for a variety of materials. Among them we mention the works by Tomiki and his collaborators⁴⁸ and by Miyata¹⁸ on alkali halides and those by Burland *et al.* on some organic molecular crystals.⁴⁹ (See also Section 10.3 for excitonic molecules.) While the general features of the observed lineshapes are in accordance with the above theory, quantitative explanation of their details seems to need more elaborate models for the scatterers (various modes of phonons and impurities) and the exciton band structures.

10.6.4 Calculated lineshapes – the low-energy tail

The semilogarithmic plots of the lineshapes (calculated with $g = 1$, $B = 0.5$ eV and $E_a = 0$ as before, and for various temperatures) in the low-energy tail part, shown in Fig. 10.14, indicate that the exponential dependence on energy of the empirical rule (10.6.1) is reproduced over a range of a few to several decades in each

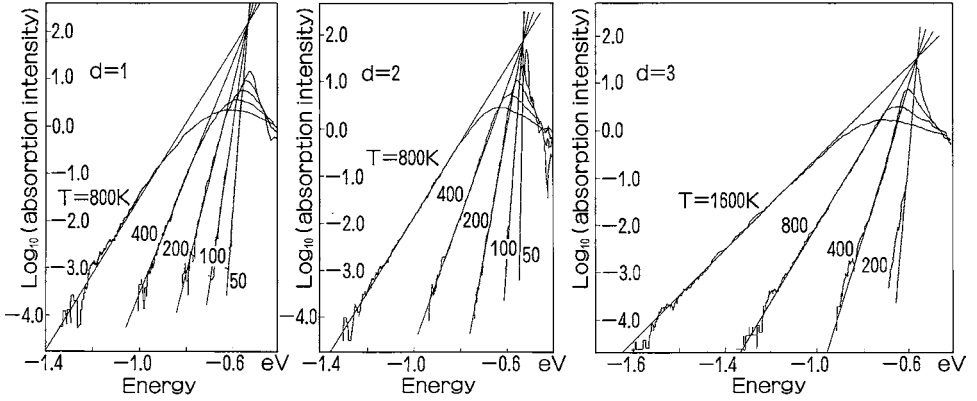


Fig. 10.14 Semilogarithmic plots of calculated absorption spectra of the *direct* exciton at various temperatures, with $g = 1$, $E_a = 0$ and $B = 0.5$ eV.³⁷

dimensionality. It is instructive to make a scaling argument before asking whether the temperature dependence is reproduced as well. If the Urbach rule (10.6.1) is derivable from our model Hamiltonian ((10.6.2) to (10.6.5)), the steepness coefficient σ in the former should be related to physical parameters appearing in the latter in the following way. Let us rescale various quantities of energy dimensionality in units of B : $(E - E_0)/B$, cQ_n/B , etc. The temperature comes in only through the Boltzmann distribution, that is, in the form $(cQ_n/B)^2 = 2E_R k_B T / B^2 = D^2 / B^2$ (see (10.6.10, 10.6.13)). Therefore, the exponent in (10.6.1) with its energy and temperature dependence can be obtained only in the form:

$$\text{const. } [(E'_0 - E)/B] / [E_R k_B T / B^2]. \quad (10.6.18)$$

This means that the steepness coefficient σ is related to the exciton–phonon coupling constant $g \equiv E_R / B$, (10.6.14), by

$$\sigma = s / g. \quad (10.6.19)$$

The dimensionless constant s , the steepness index, no longer depends on the material constants (such as g) but only on the geometrical structure of the lattice, in particular on its dimensionality.

Coming back to the temperature dependence of the decay constants in Fig. 10.14, one finds proportionality to T^{-1} for $d = 2$ and 3 in conformity with (10.6.18) but to $T^{-2/3}$ for $d = 1$. This is shown by the *broken* lines in Fig. 10.15, where the value $s = \sigma \times g$ obtained by best fitting of (10.6.1) to the calculated lineshape at each temperature is plotted by the symbols \circ , \triangle and \square for $d = 1, 2$ and 3 .

By inverting the sign of the transfer energy t , one obtains another model for band structure with the direct exciton peak ($K = 0$) at the top and the *indirect* exciton edge at the bottom of the band. Figure 10.16 is a semilogarithmic plot

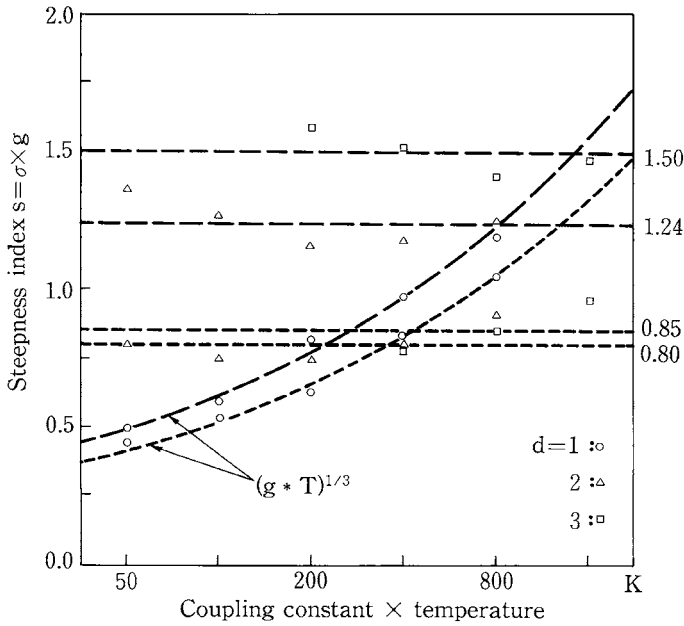


Fig. 10.15 Calculated steepness index s (plotted by \circ , \triangle and \square for $d = 1, 2$ and 3 , respectively) of the Urbach tail for direct and indirect excitons (longer- and shorter-dashed lines, respectively). The one-dimensional results are best fitted by $T^{1/3}$ lines.³⁷

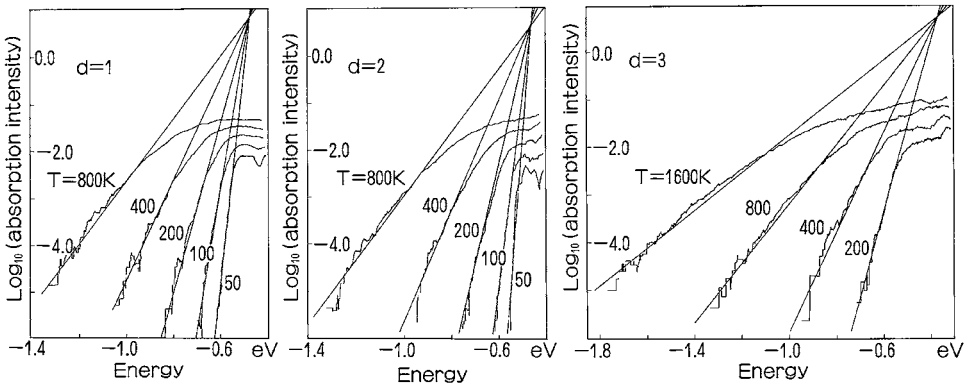


Fig. 10.16 Semilogarithmic plots of calculated absorption spectra of the *indirect* exciton at various temperatures, with $g = 1$, $E_a = 0$ and $B = 0.5$ eV.³⁷

of the thus-calculated indirect absorption edge, with the same parameter values as before. One finds again an exponential dependence on energy over a range of 2.5–4 decades of intensity for various temperatures. The temperature dependence of the decay constant is the same as before as shown by the short-dashed lines in Fig. 10.15. The slope is more gentle in the indirect edge than in the direct edge for each dimensionality.

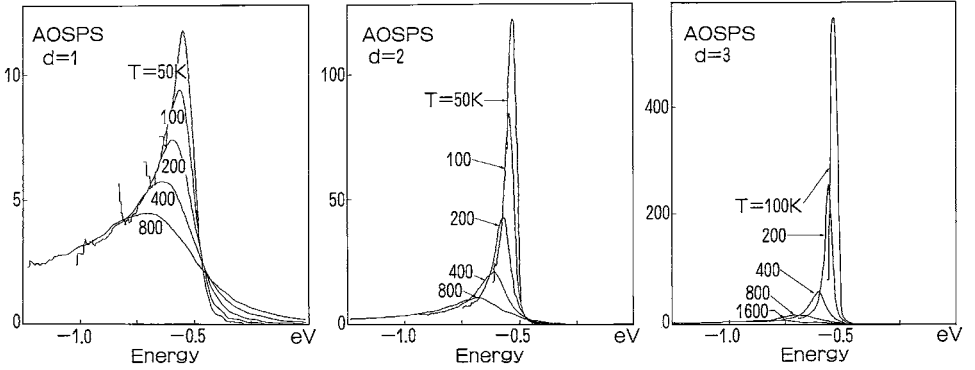


Fig. 10.17 Calculated AOSPS (average oscillator strength per state) spectra, with the same parameter values as in Fig. 10.14 and 10.16.³⁷

In order to see the physical origins of the low- and high-energy tails of the absorption spectra, we show in Fig. 10.17 the calculated AOSPS spectra $f(E)$ defined by (10.6.17), for the direct edge only. (The calculated density-of-states curves are not shown here.) What is most remarkable here is the *energy and temperature dependence which is completely different on the high- and low-energy sides of the peak*. For all dimensionalities, the high-energy side of $f(E)$ fits to a Lorentzian (with temperature-dependent width as shown in Fig. 10.13) much better than $F_a(E)$ does. This indicates that the high-energy tail is due to the scattering-limited lifetime broadening of a delocalized exciton, as was already discussed in Subsection 10.6.3. In contrast, the low-energy tail is rather insensitive to temperature and decays with decreasing energy more rapidly for higher dimensionality. These features clearly indicate that the low-energy tail is due to the momentarily localized exciton states as has been postulated.^{15,39,41} As shown in (8.8.10) for the case of a single-site attractive potential and by subsequent qualitative analysis applicable to more general cases, the oscillator strength of a localized state is essentially a function of its binding energy only (measured from the effective band edge \bar{E}_0 which is close to the absorption peak E_p), being insensitive to the momentary potential, $V_n = -cQ_n$, of the individual site. In fact, one finds that the low-energy tail of $f(E)$ in Fig. 10.17 decays as $(E_p - E)^{-d/2}$ in accordance with (8.8.12) if one takes for E_p the peak of the calculated $F_a(E)$ at each temperature.

Starting from the model Hamiltonian ((10.6.2) to (10.6.5)), one can thus derive numerically the empirical rule (10.6.1) without being prepossessed by any particular physical picture and without the use of any approximation. Through the analysis of the calculated spectra of $f(E)$, however, we have been able to unambiguously ascribe the exponential tail of $F_a(E)$ to the momentarily localized exciton states. Thus, the introduction of ATA and CPA into the present problem turns out to be an important step towards reality; the remaining difficulty inherent to these one-site

approximations, however, has to be resolved by resorting to direct numerical calculation.

A question may well arise here. Can theoreticians be really satisfied with the *numerical derivation* of an *analytically very simple empirical rule*? The answer is that the former is the only way of directly connecting the proposed model Hamiltonian to the latter without any approximation whose accuracy is difficult to assess. A useful hint concerning this question, if not an answer, may be found in the scaling argument given for this problem by Ihm and Phillips.⁵⁰ As another question, has the two-mode model nothing to do with the present numerical theory? The former gave an analytical explanation of the Urbach rule for an impurity-bound exciton, and one is tempted to explain the apparently same rule for intrinsic and extrinsic excitons by the same mechanism. It should be mentioned, in this connection, that the numerical calculation of lineshapes for a crystal with an impurity was also performed using the same Hamiltonian as above (except for the inclusion of an attractive potential at the impurity site), and that the Urbach tails were thereby obtained for extrinsic (the last paper of Ref. [37]) as well as for intrinsic edges. It is easy to show that the numerical theory is equivalent to the two-mode model as far as the extrinsic absorption band (discrete level) is concerned. It seems difficult, however, to find out how and where the two-mode mechanism is concealed in the numerical theory for the intrinsic exciton.

10.6.5 Comparison with experiments

The relation (10.6.19) between σ and g , together with the values of s obtained by the numerical calculation, serves to *correlate two independent groups of experiments*. While σ is obtained from the absorption spectra, g characterizes the emission spectra which are predicted to be F-type or S-type according as $g < \text{or} > g_c = 1 - (2z)^{-1}$ (Subsection 10.6.2, Fig. 10.11). Considering a simple cubic lattice for which $g_c = 0.917$ and $s = 1.50$ (direct edge) or 0.85 (indirect edge), we find the critical values $\sigma_c = 1.64$ (direct) or 0.93 (indirect) below which the exciton is expected to be self-trapped.

Table 10.1 shows some examples of materials for which the Urbach rule has been established for the direct or indirect edge, in increasing order of σ . In the column headed “emission” we indicate the nature of relaxed excitons distinguished by the emission spectra – the Stokes-shifted broad emission band from a self-trapped exciton (S) or the resonant sharp emission line from a nearly free exciton (F). Thus, the two groups of experimental data can be correlated with each other in satisfactory consistence with the theoretical predictions on σ_c in spite of the simplified model used for that.

Table 10.1 *Correlation between the observed steepness coefficient σ of the absorption spectra and the observed nature (F or S) of the relaxed exciton judged from the emission spectra in various crystalline materials (see Sections 10.6, 9.5 and 9.6).*

Direct exciton (predicted $\sigma_c = 1.64$)					Indirect exciton (predicted $\sigma_c = 0.93$)			
Material	Reference	Observed σ	Emission		Material	Reference	Observed σ	Emission
α -Al ₂ O ₃	[51]	0.505 ($E\parallel c$) 0.672 ($E\perp c$)	?					
Se	[52]	0.6	?					
TiO ₂ (anatase)	[53]	0.621 ($E\parallel c$) 0.733 ($E\perp c$)	S					
SiO ₂	[54]	0.63						
SnO ₂	[55]	0.7 ^a						
NaCl	[57]	0.76	S					
KBr	[58]	0.79	S					
KCl	[59]	0.80	S					
KI	[36]	0.82	S					
α -perylene	[60]	0.93	S					
β -perylene	[60]	1.38	S					
GeS	[61]	1.45	?					
PbI ₂	[62]	1.48	S \sim F		AgCl	[68]	0.78	S
anthracene	[63]	1.5–1.7	F		AgCl _{0.45} Br _{0.55}	↓	0.89	S = F
CdS	[64]	2.2	F		AgBr	[68]	0.98	F
ZnSe	[65]	2.4	F		TiO ₂ (rutile)	[53]	0.90	F
CdSe	[66]	2.5	F		TlCl	[69]	1.16	F
ZnTe	[65]	2.8	F					
CdTe	[66]	3.0	F					
Te	[67]	3.1	F					

^a Direct forbidden edge, see Ref. [56].

Of particular interest in Table 10.1 is the indirect exciton in $\text{AgBr}_{1-x}\text{Cl}_x$, of which the gradual decrease of σ with increasing x (see Fig. 10.18) and the abrupt F–S transition of the luminescence (see Fig. 10.19) were observed and correlated by Kanzaki, Sakuragi and Sakamoto.⁶⁸ Moreover, the value of σ they found at the F–S

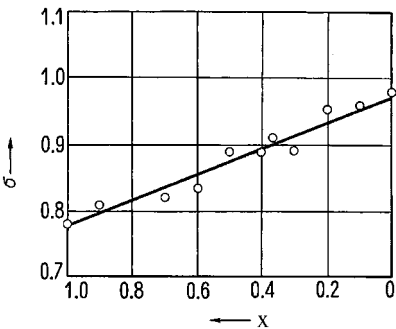


Fig. 10.18 Observed change of steepness coefficient σ in the Urbach rule for the absorption edge (*indirect*) of $\text{AgBr}_{1-x}\text{Cl}_x$ at 293 K against composition x . Due to Kanzaki *et al.*⁶⁸

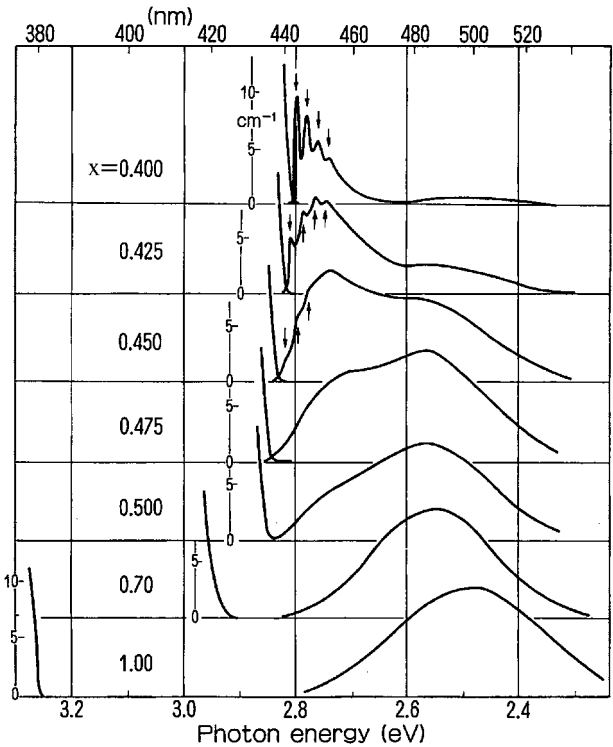


Fig. 10.19 Absorption edge (*left*) and photoluminescence (*right*) spectra at 2 K of $\text{AgBr}_{1-x}\text{Cl}_x$ for $x \geq 0.4$. Due to Kanzaki *et al.*⁶⁸

transition turned out to be in excellent agreement with σ_c obtained by numerical calculations.

Another interesting fact is the observation of the Urbach rule in the emission⁷⁰ as well as in the absorption spectra of the exciton in PbI_2 , the former obeying

$$F_c(E) \propto \exp[-(\sigma - 1)(E_0^i - E)/k_B T] \quad (10.6.20)$$

where σ is the steepness coefficient which appeared in the Urbach rule for the absorption spectra of the same material.⁶⁰ Namely, the absorption and emission spectra in this narrow spectral region of the exponential tail satisfy the general relation (3.2.3) derived from the Einstein relation under the assumption of thermal equilibrium in the excited state. The fact that the exciton in this material ($\sigma = 1.48$) is close to the (theoretical) F-S boundary ($\sigma_c = 1.64$) may be helping to bring it into thermal or quasi-thermal equilibrium before emitting a photon.

A similar situation may exist in pyrene ($\sigma = 1.38$),⁷¹ β -perylene ($\sigma = 1.38$) and even in α -perylene ($\sigma = 0.93$)⁷² the emission spectra of these materials consist of two components originating from F- and S-excitons in quasi-thermal equilibrium with an F-S energy difference of the order of one to several hundred cm^{-1} . This is in contrast to the alkali halides in which F- and S-excitons are far from thermal equilibrium (see Section 9.5). In aromatic compounds, B is small and hence the F-S energy difference is small even if σ is not so close to σ_c . In connection with this, some aromatic crystals with σ close to σ_c , such as anthracene and tetracene can be brought from the F-type to the S-type by applying high pressure.⁷³

In contrast, some of II-VI compound semiconductors have large σ -values, being consistent with the weak exciton-phonon coupling constant generally accepted in these materials. At low temperatures, however, the low-energy tail of the exciton absorption deviates significantly from the exponential form, and a perturbation-theory approach to describe the tail as the sidebands of multiphonon absorption was proposed by Segall⁷⁴ in analysing the measurement on CdTe due to Marple.⁷⁵ This is not inconsistent with our semi-classical approach which is expected to give the high-temperature asymptote of such sidebands. It is no wonder that some structures due to successive phonon sidebands reveal themselves in weak electron-phonon-coupling materials. In fact, Liebler *et al.*⁷⁶ showed by cumulant expansion calculation of the lineshape that the multi-optical-phonon sidebands approach the Urbach lineshape at high temperatures.

In connection with this, it is interesting to note the different behaviors of TiO_2 depending upon its crystal structure; the Urbach rule has been observed and correlated with the emission spectra by Tang, Levy, Berger and Schmid.⁵³ The anatase form shows Urbach tails with $\sigma \sim 0.6$ – 0.7 down to the lowest temperature measured, in contrast to the rutile form whose Urbach tail with $\sigma \sim 0.9$ changes to step-like phonon sidebands at low temperatures (see Table 10.1). In the former,

the S-type luminescence had in fact been found and ascribed to the self-trapped exciton,⁷⁷ “which was later confirmed to be polarized as $E \perp c$,⁷⁸ the same as the direct-exciton Urbach tail situated lower than the indirect-exciton tail with $E \parallel c$. According to a recent study by time-resolved spectroscopy,⁷⁹ this S-type emission consists of two components. The exponential decay component which is dominant for the excitation energy near the absorption edge is ascribed to the direct process of exciton self-trapping. The power-law decay component which increases for the higher excitation energy is supposed to originate from once ionized electron-hole pairs. In contrast to anatase, a free-exciton luminescence was observed in purified rutile⁸⁰ which is consistent with the above-mentioned behavior of the absorption edge at low temperatures. The value of $\sigma = 0.9$ is not inconsistent with $\sigma_c = 0.93$ for the indirect edge; note that this Urbach tail is associated with a weak indirect edge, not with the strictly forbidden direct edge.

The fact that the weakly phonon-coupled excitons without self-trapped states reveal step-like structures due to phonon absorption at low temperatures even if they obey the Urbach rule at high temperatures is consistent with the renormalization theory in Section 10.2 applied to the zero-phonon line and the phonon sidebands in Section 10.4. The renormalization theory is convergent in the absence of self-trapping, and as T tends to zero the phonon-absorbing sidebands (hot bands) should disappear, making the absorption edge very sharp even if the zero-phonon line cannot be separated from the phonon-emitting sidebands. Admittedly, such an idealized situation may be difficult to realize since real crystals contain some amounts of lattice imperfections giving rise to shallow bound excitons which contribute giant oscillator strengths (see Section 8.8 and Subsection 10.6.4) to the low-energy tail of the absorption.

In contrast, it is natural to expect that strong exciton-phonon coupling favors the adiabatic approach and smooth lineshape even at low temperatures although $k_B T$ in the empirical Urbach rule must be replaced by the effective zero-point energy $\hbar\omega/2$ at the lowest temperature. These points will be studied more carefully in the following section.

10.7 Examination of model Hamiltonians

The questions remain as to whether and to what extent the simplified model Hamiltonian (10.6.2 to 10.6.5) represents realistic situations. For instance, (i) does the localized mode Q_n represent the acoustic mode or the optical mode, or both? (ii) Can the interaction between a large-radius exciton and the lattice vibrations be described by the short-range and on-site interaction (10.6.4)? We will begin by giving three model systems to which the Hamiltonian is applicable.

The first is an exciton, with its internal structure *ignorable* due to a large binding energy, which is subject to the deformation potential of the acoustic-phonon field. Approximate the exciton kinetic energy by

$$K_L = (c_\ell^2/2a_0^2) \sum_n \sum_{n'} \sum_l [\delta_{n,n'} - \delta_{n,n'-l}] P_n P_{n'}. \quad (10.7.1)$$

Here the vector l runs over six nearest neighbors $(\pm 1, 0, 0)$, $(0, \pm 1, 0)$ and $(0, 0, \pm 1)$ of the simple cubic lattice. The Fourier transform

$$P_n = N^{-1/2} \sum_k \exp(-ik \cdot na_0) P_k, \quad Q_n = N^{-1/2} \sum_k \exp(-ik \cdot na_0) Q_k \quad (10.7.2)$$

gives in fact

$$K_L = (1/2) \sum_k \omega_k^2 P_k^* P_k, \quad U_L = (1/2) \sum_k Q_k^* Q_k, \quad (10.7.3)$$

$$\omega_k^2 = (c_\ell^2/a_0^2) \sum_{xyz} 2[1 - \cos(k_x a_0)] \cong (c_\ell k)^2 \quad (ka_0 \ll 1), \quad (10.7.4)$$

as is appropriate for the acoustic mode with longitudinal sound velocity c_ℓ . According to (9.4.1, 9.4.3) the deformation potential for a structureless exciton is given by the diagonal Hamiltonian H_I with

$$\langle n | H_I | n \rangle = \Xi \Delta_n = -c Q_n, \quad (10.7.5)$$

while the elastic energy of the lattice is given by

$$U_L = (Ca_0^3/2) \sum_n \Delta_n^2, \quad (10.7.6)$$

with C being an appropriate elastic constant. Equations (10.7.1., 10.7.5, 10.7.6) are in accordance with our model Hamiltonian, (10.6.4, 10.6.5). From (10.7.2, 10.7.4, 10.7.5) one finds

$$c = a_0^{-3/2} C^{-1/2} \Xi. \quad (10.7.7)$$

The second and obvious example is an exciton with short-range interaction with the optical mode with dispersionless frequencies: $(\omega^2)_{nn'} = \delta_{nn'} \omega_0^2$ (the so-called Einstein model) which will be considered later in more detail. Since the site number n as it appears in the interaction $\langle n | H_I | n \rangle$ denotes the center-of-mass coordinate of the exciton, this on-site interaction is not subject to the neutrality cancellation as far as $m_e \neq m_h$.

Thirdly, under the coexistence of the above two modes which, together with their interaction coefficients c , shall be suffixed by ac and op, respectively, one can define the interaction mode Q_n by

$$\langle n | H_I | n \rangle = -c Q_n \equiv -c_{ac} Q_n^{ac} - c_{op} Q_n^{op} \quad (10.7.8)$$

and the *non-interacting mode* Q_n' after an orthogonal transformation $(Q_n^{ac}, Q_n^{op}) \rightarrow (Q_n, Q_n')$, and rewrite the Hamiltonian with the interaction modes

alone, leaving the non-interacting modes aside since the latter have no effect on the lineshape (remember how the interaction mode was introduced in Section 4.4). This is our answer to question (i). The acoustic mode ubiquitous in all crystals has the effect of smoothing the lineshape because of its continuous spectrum. For this reason, the classical approximation for the interaction mode Q_n will provide a useful description of the overall lineshape, especially at high temperatures where the structures due to the optical phonons are blurred by the participation of acoustic modes.

The internal structure of an exciton (electron–hole relative motion), which is incorporated into the scattering matrix element (10.1.1) with (10.1.2) to (10.2.4), retains the site diagonality of H_I , but its diagonal element $\langle n|H_I|n \rangle$ now contains $Q_{n'}$ of the neighboring sites n' which are within the distance of the exciton radius, though with smaller coefficients than c . They could be incorporated in an extended version of the model Hamiltonian, with which the numerical calculation could be performed to see whether the exponential rule still holds or not. Such an extended calculation could give an answer to question (ii).

Apart from the Urbach rule, it is instructive to complement the semi-classical approach to realistic systems in the preceding section by a quantum-mechanical study of an idealized model Hamiltonian. A typical one is the Einstein version of (10.6.2 to 10.2.5) with $(\omega^2)_{nn'} = \omega^2 \delta_{nn'}$ and with nearest neighbor transfer $t(<0)$. Taking the origin of energy at E_a (atomic or molecular excitation energy) $+ N\hbar\omega/2$, one can write

$$H = \sum_m \sum_{\ell} |m\rangle t \langle m + \ell| - (E_R \hbar \omega)^{1/2} \sum_m |m\rangle (b_m + b_m^\dagger) \langle m| + \hbar \omega \sum_m b_m^\dagger b_m, \quad (10.7.9)$$

where m extends over all sites, ℓ over z nearest neighbor sites, and b_m, b_m^\dagger are, respectively, the annihilation, creation operators of a phonon (see (2.2.3, 2.2.4)) at site m . The first term of (10.7.9) gives the unperturbed exciton band with energy given by

$$t(\mathbf{K}) = t \sum_{\ell} \exp(i \mathbf{K} \cdot \mathbf{R}_{\ell}). \quad (10.7.10)$$

Here we have three physically independent parameters: the half-width $B = z|t|$ of the exciton band, the relaxation energy E_R and the phonon energy $\hbar\omega$. The dimensionless parameter, $S \equiv E_R/\hbar\omega$, is called the electron–phonon coupling strength. This model Hamiltonian has been studied theoretically by a number of authors in the context of a variety of problems.^{81–86}

In the wide-band case, $B \gg \hbar\omega$, the stable state of an exciton is classified rather distinctly to a nearly free state and a self-trapped state according as $g \equiv E_R/B < \text{or} > g_c(\sim 1)$ as was studied in detail with the use of various model

Hamiltonians in Sections 9.4 to 9.6. However, the distinction becomes obscure as the ratio $B/\hbar\omega$ decreases down to 3 or 4.

In the opposite case of a narrow band, $B < \hbar\omega$, as is realized in molecular crystals with small intermolecular transfer energy and large intramolecular vibrational quanta (see Section 10.5), it is more appropriate to start from the n th vibrational state (explicit form given by (2.2.17)) of a molecule with energy $n\hbar\omega$ and then consider the vibronic band formed therewith:

$$|n, \mathbf{K}\rangle = N^{-1/2} \sum_m \exp[i\mathbf{K} \cdot \mathbf{R}_m] (n!)^{-1/2} (b_m^\dagger)^n |0\rangle_{\text{ph}} |m\rangle, \quad (10.7.11)$$

$$E_n(\mathbf{K}) = n\hbar\omega + I_n t(\mathbf{K}), \quad (10.7.12)$$

$$I_n = \exp(-S) S^n / n!, \quad (10.7.13)$$

where $|0\rangle_{\text{ph}}$ denotes the phonon vacuum state. In order that the n th vibronic exciton is transferred from site m to its neighbor $m + 1$, the molecule on site m must return from the n th to the 0th vibrational states while that on the neighbor from the 0th to the n th. Hence, one has the reduction factor I_n for the transfer energy t where I_n is the square of the overlap integral between the 0th and n th vibrational states (see eqs. (4.5.5, 4.5.6)).

Let us consider the 0th and 1st vibronic bands (0) and (1) which are extended on the energy intervals $\pm B \exp(-S)$ and $\hbar\omega \pm BS \exp(-S)$, respectively, and the continuum (0') of the former plus one-phonon extended on the interval $\hbar\omega \pm B \exp(-S)$. The band (0) does not overlap the continuum (0') as long as $2B/\hbar\omega < \exp(S)$, while the band (1) stays within or just out of the continuum (0') according as $S < \text{or} > 1$. Thus we have different types of energy diagram depending on the parameter values S and $B/\hbar\omega$. The dispersion of the vibronic bands, defined as the functional dependence of $E_n(\mathbf{K})$ on $t(\mathbf{K})$, will be subject to change when the interaction between different vibronic bands, so far neglected, is taken into account.

Sumi⁸⁶ presented an extensive study of this model Hamiltonian over the entire range of dimensionless parameters S and B ($\equiv B/\hbar\omega$ of the above notation), applying the *dynamical coherent potential approximation* (d-CPA), devised by himself as an extension of the CPA (see Section 8.7) to the case of *temporally fluctuating potential*. Some of the results he obtained will be reproduced here. In Fig. 10.20 are shown some examples of the energy (in unit of $\hbar\omega$) diagram of vibronic bands (solid lines) and their plus-one-phonon continuums (hatched regions) as functions of the bare band energy $t(\mathbf{K})$ (in units of $2B$) for different sets of parameter values. Note that the bottom of the lowest vibronic band is taken as the origin of energy. One finds here not only the varied features of the energy diagram mentioned above but also the abrupt increase of the effective mass of the exciton polaron (the reciprocal

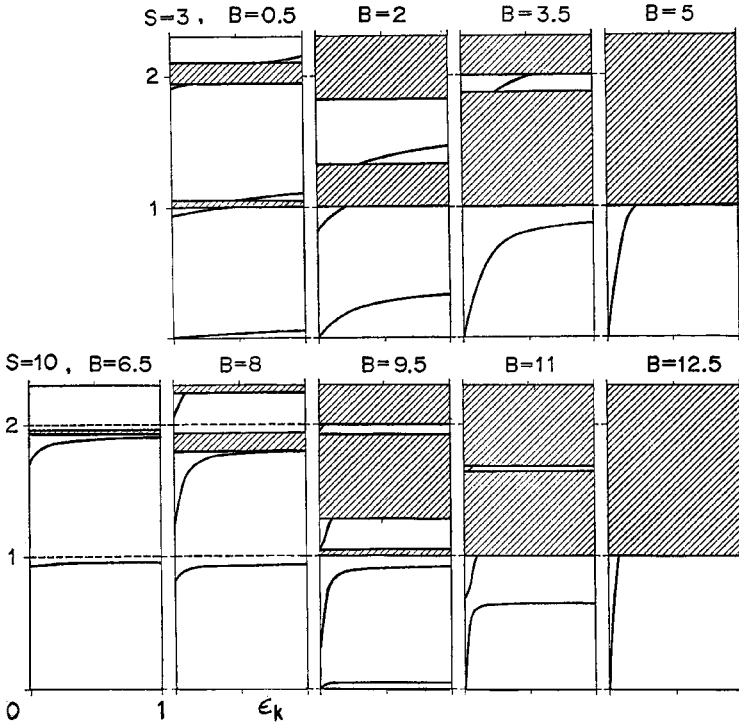


Fig. 10.20 Energy diagram of the exciton-phonon system for $S = 3$ (upper half) and $S = 10$ (lower half). Note that $(E_k - E_0)/\hbar\omega$ on the ordinate is plotted as a function of $(t_k - t_0)/2B$. Due to Sumi.⁸⁶

of the initial slope of the lowest vibronic band) as B becomes less than about S , especially when S is large (as is expected for wide band, $B \gg \hbar\omega$). Some of the calculated exciton absorption spectra are shown in Fig. 10.21. One finds how the zero-phonon line (at the bottom of the lowest vibronic band, chosen as the origin of energy) splits from the main absorption peak as B becomes less than about S , and soon becomes unobservably weak indicating that the exciton is self-trapped. If one looks at the *envelope* of the absorption peaks, one clearly observes its motional narrowing as B increases against S , which reminds us of Fig. 10.2 indicating the close relationship between CPA and d-CPA.

10.8 Polariton bottleneck for exciton-photon conversion

As we have seen in the preceding section, the features of the emission spectra are quite different according as the exciton as the initial state is in the nearly free state (F) or the self-trapped state (S). Confining ourselves to the former case, we note that there is another difference in the emission process and hence in the emission

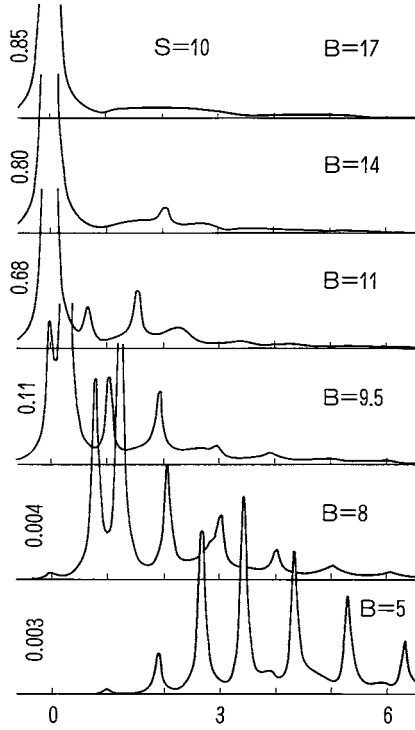


Fig. 10.21 Absorption spectra for various values of $B (\equiv B/\hbar\omega$ of the text) with fixed $S = 10$. The zero of energy and the unit of the abscissa are taken to be the lowest exciton–polaron energy and $\hbar\omega$, respectively. Due to Sumi.⁸⁶

spectra between the direct and the indirect exciton. In the indirect case, an exciton thermalized (in a time of the order of 10^{-12} – 10^{-11} s at the band bottom at $\mathbf{K}_b \neq 0$) emits a photon (in a time of the order of 10^{-6} s with simultaneous emission or absorption of a phonon – the indirect radiative annihilation which can be described by the second-order perturbation theory (the inverse of the indirect radiative creation described in Section 10.2). The width of each emission line corresponding to each participating mode is expected to be of the order of $k_B T$ reflecting the Boltzmann distribution of the excitons.

The situation is not only different but even embarrassing in the direct case. If one were to consider this as the limiting case, $\mathbf{K}_b \rightarrow \mathbf{0}$, of the indirect transition mentioned above, the energy denominator $E_0 \pm \hbar\omega_{\pm\mathbf{K}_b} - E_{\mathbf{K}_b}$ in the second-order matrix element would tend to zero as $\mathbf{K}_b \rightarrow \mathbf{0}$ in the case of acoustic mode, resulting in the divergence of the emission rate. If one considers the problem straightforwardly assuming the excitons thermalized around the band bottom at $\mathbf{K} = 0$, the square of the optical transition matrix element becomes of the order of $N \sim \infty$ singularly only at $K \simeq 0$, invalidating the first-order perturbation theory.

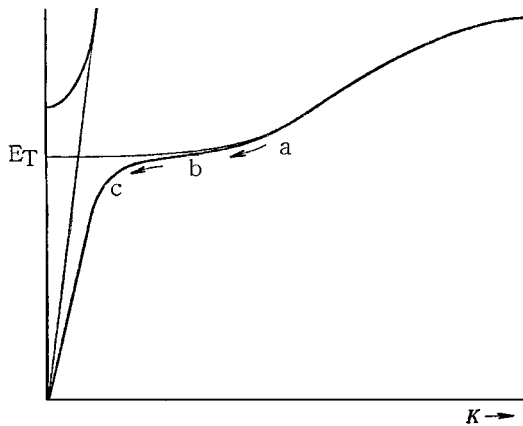


Fig. 10.22 Polariton picture for direct radiative annihilation of an exciton in the weak coupling case, with bottleneck between b and c.⁸⁷

This is the very region where the polariton effect is essential as was described in Section 8.6.

According to the polariton picture which is reproduced in Fig. 10.22, the conversion of an exciton into a photon takes place gradually,⁸⁷ not as a quantum transition process as is described by first- or second-order perturbation theory. An exciton (or electron-hole pair), excited by light well above the lowest exciton absorption peak, will tend to be thermalized, first by emitting optical phonons (the most efficient way of losing kinetic energy as far as they can do so), and then by emitting or absorbing acoustic phonons, towards the band bottom ($a \rightarrow b$ of Fig. 10.22). If the exciton can somehow reach the region c of the polariton with a significant mixture of a photon, it has a good chance of getting out of the crystal as a photon so as to be observed as emission, as is also obvious from the reflectivity curve (c) of Fig. 8.6 according to which a polariton with $\omega < \omega_t$ has a chance of not being totally reflected ($R(\omega) < 1$) at the crystal surface. There is a bottleneck, however, for the polariton b near the band bottom impeding its reaching region c for the following reason. The matrix element as well as the density of final states for the acoustic-phonon scattering decrease as the exciton wave vector \mathbf{K} decreases, resulting in the retardation of thermalization as one approaches the band bottom. The conversion rate of the polariton into an external photon increases due to the decrease in $R(\omega)$ with the decrease of polariton energy $\hbar\omega$ as mentioned above. As a result, the polaritons tend to be populated within a more or less limited region of $K \equiv |\mathbf{K}|$, called a “bottleneck”, forming a quasi-steady state under the balance of inflow from the higher- K region of photoexcitation through energy loss by phonon scattering and the outflow to the lower- K region where they are subject to a higher probability of conversion to external photons.

The earliest experiment to confirm the existence of a bottleneck was performed by Heim and Wiesner with the use of time-resolved spectroscopy of photoluminescence.⁸⁸ They found the decay time of the polariton in CdS to depend on its energy, indicating the non-equilibrium population at low temperatures, and could locate the bottleneck at ~ 1 meV below the transverse exciton energy $E_t (= E_0$, the band bottom mentioned above) on the basis of various experimental facts, and in particular by showing that the decay time increases as the polariton energy decreases towards this region.

Sumi⁸⁹ presented a quantitative theory of this bottleneck effect by taking into account the statistical distribution of polaritons and their scattering by optical as well as acoustic phonons and the impurity quenching of luminescence. Thereby he calculated the lineshape of the polariton luminescence (zero-phonon line) and its optical phonon (emission) sideband in CdS, and succeeded in explaining the observed width of the former as well as the intensity ratio of the latter to the former which are both much larger than predicted by the non-polariton thermal equilibrium theory. Figure 10.23 is the semilogarithmic plot of calculated (a) radiative decay rate and (b) phonon-assisted decay rate as functions of polariton energy around the exciton band bottom E_0 . Their sum makes a deep valley at slightly below E_0 , a clear indication of the bottleneck.

Askary and Yu⁹⁰ studied the decay of the photoluminescence in CdS and CdSe by time-resolved spectroscopy and analysed the results elaborately by taking into

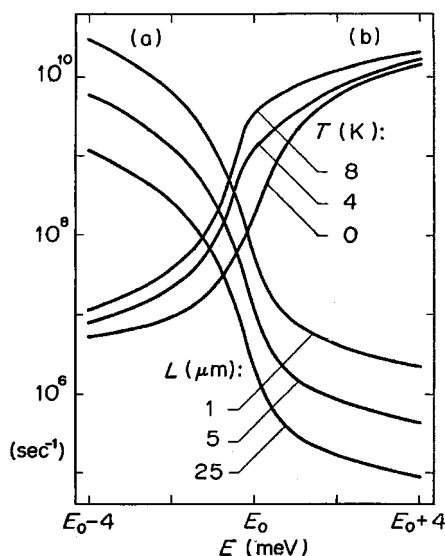


Fig. 10.23 (a) Calculated radiative decay rate for the crystal thickness $L = 1, 5$ and $25 \mu\text{m}$ and (b) phonon-assisted decay rate for temperature $0, 4$ and 8 K , of a lower-branch polariton in CdS with energy around the lowest exciton energy $E_0 = E_t$. Due to Sumi.⁸⁹

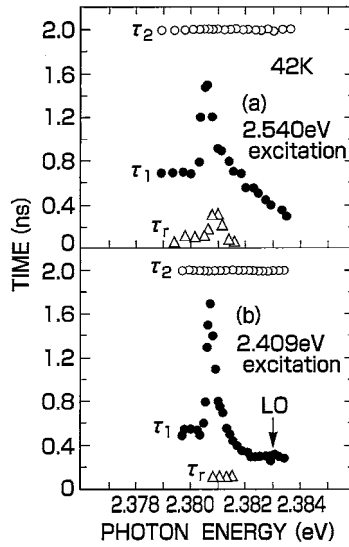


Fig. 10.24 Energy dependence of the decay times τ_1 , τ_2 and the rise time τ_r of polaritons in ZnTe deduced by the deconvolution calculation. Due to Oka *et al.*⁹¹

account the existence of the upper-branch polariton (which in fact shows up as a kink in the luminescence spectra) with the use of Pekar's additional boundary condition, and confirmed that the decay time has in fact a *maximum* at the bottleneck approximately the same as was located by Heim and Wiesner.

Oka, Nakamura and Fujisaki⁹¹ studied this problem in ZnTe by picosecond time-resolved spectroscopy, in the context of elucidating the evolution of Raman scattering via hot to ordinary photoluminescence. As shown in Fig. 10.24, they found two exponential decay components of the photoluminescence with time constants τ_1 , τ_2 , and one decay component τ_r on the supply side, as functions of the emitted photon energy E , similar to the result of Askary and Yu on CdS. The longest component $\tau_2 = 2$ ns which is E independent and is equal to the decay time of the time-integrated intensity represents the lifetime of quasi-thermalized polaritons, while τ_1 shows a sharp maximum of 1.5 ns at E a little below the transverse exciton energy E_0 , indicating the temporal change of polariton redistribution piling up towards the bottleneck. The shortest component τ_r represents the early rise time of the polariton population in the bottleneck region supplied from the higher-energy region of optical excitation. They found furthermore the effective temperature of the distribution of polaritons well after reaching the steady state (at 1 ns after photoexcitation) to be 16 K, which is considerably higher than the lattice temperature 4.2 K, whereby they emphasized that the polariton population is still “hot” even in the steady state.

11

Higher-order optical processes

11.1 Nonlinear responses and multiphoton processes

The invention of the laser – a light source device making use of “light amplification by stimulated emission of radiation” as its working principle¹ – brought about revolutionary developments in the spectroscopic study of matter. The *high intensity, monochromaticity, coherence* and *directivity* of this new light source,² have all contributed to the development of adaptable technologies which can meet a variety of the high-grade requirements in spectroscopic research. In particular, the intensity can be made high enough to give rise to *nonlinear responses* of matter up to quite high orders.³ Fortunately, these responses can be well separated into successive orders, on the basis of their dependence on the amplitude of the incident electromagnetic field and their characteristic frequency dependence, unless the intensity is too high. In fact, the response can be expanded in a power series in the field, the good convergence of which originates primarily from the *smallness of the radiation–matter coupling constant* compared to other interactions within the matter.

To make the statement more definite, we start with the lowest-order nonlinear response. Consider two independent electromagnetic waves, $\mathbf{E}_1 \exp(i\mathbf{k}_1 \cdot \mathbf{r} - i\omega_1 t)$ and $\mathbf{E}_2 \exp(i\mathbf{k}_2 \cdot \mathbf{r} - i\omega_2 t)$ propagating in the same direction ($\mathbf{k}_1 \parallel \mathbf{k}_2$) within matter with refractive index $n(\omega)$ as defined by (1.2.5). Ignoring their spatial dependence for the moment, one can consider that part of the polarization which is induced as a second-order response to \mathbf{E}_1 and \mathbf{E}_2 of the form (cross terms)

$$\mathbf{P}^{(2)} = \chi^{(2)}(\omega; \omega_1, \omega_2) : \mathbf{E}_1 \mathbf{E}_2 \quad (\omega = \omega_1 + \omega_2), \quad (11.1.1)$$

corresponding to the linear response given by (6.1.29). Here the second-order susceptibility $\chi^{(2)}$ represents a tensor of the third rank with components $\chi_{ijk}^{(2)}$, the coefficients connecting the i th component of \mathbf{P} to the j th and k th components of \mathbf{E}_1 and \mathbf{E}_2 , respectively. In the same way as was done in deriving (6.1.31) for the

linear susceptibility χ , one can derive the microscopic expression for the second-order susceptibility $\chi^{(2)}$ which consists of terms with the products of three (instead of two) matrix elements of the P s and two (instead of one) energy denominators corresponding to two intermediate states, with the use of second-order perturbation theory once using the $-\mathbf{P} \cdot \mathbf{E}_1$ term and once using the $-\mathbf{P} \cdot \mathbf{E}_2$ term of the perturbation Hamiltonian (see the second term on the r.h.s. of (6.1.23)). It is obvious from this derivation that the resulting polarization $\mathbf{P}_\omega^{(2)}$ has angular frequency $\omega = \omega_1 + \omega_2$, and wave vector $\mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2$ if the spatial factors are revived in the derivation.

This $\mathbf{P}_\omega^{(2)}$, as a new source, will generate a new electromagnetic wave $\mathbf{E}_\omega^{(2)}$ with the same ω and \mathbf{k} through the equation

$$\nabla^2 \mathbf{E}_\omega + [n(\omega)^2 \omega^2 / c_0^2] \mathbf{E}_\omega = - [\omega^2 / \epsilon_0 c_0^2] \mathbf{P}_\omega^{(2)} \quad (11.1.2)$$

obtained from eqs. (1.1.1, 1.12) with $J \equiv \partial \mathbf{P} / \partial t$ as the source term and with the use of the refractive index $n(\omega) \equiv c_0 / c = c_0 k / \omega$ (see the definition (1.2.5)), provided, in the case of parallel \mathbf{k} s, that the *phase-matching condition*

$$k = n(\omega)\omega / c_0 = n(\omega_1)\omega_1 / c_0 + n(\omega_2)\omega_2 / c_0 \quad (\omega = \omega_1 + \omega_2) \quad (11.1.3)$$

is fulfilled. In a trivial case of a vacuum with $n(\omega) = 1$ where it holds identically, we have unfortunately no optical nonlinearity. In isotropic matter, it is possible neither within a region of normal dispersion ($n(\omega)$ increasing, hence $n(\omega) > n(\omega_1), n(\omega_2)$) nor within that of anomalous dispersion ($n(\omega)$ decreasing, hence $n(\omega) < n(\omega_1), n(\omega_2)$). It can be fulfilled however in anisotropic matter if one makes use of the different polarization directions of \mathbf{E}_1 and \mathbf{E}_2 with different $n(\omega)$ functions. Under the phase-matching condition satisfied in this way, the electromagnetic wave with sum frequency ω will spatially grow as it proceeds in the same direction as the incident waves with ω_1 and ω_2 .

In this way, the electromagnetic wave with *sum frequency* $\omega = \omega_1 + \omega_2$, and naturally the *second harmonic* wave with $\omega = 2\omega_1$, too, can be generated from those with lower frequencies ω_1 and ω_2 through the second-order susceptibility $\chi^{(2)}$, thus contributing to obtaining a light source with higher frequency via the use of lasers with lower frequencies. In the same way, one can also generate the electromagnetic wave with *difference frequency*, $\omega = \omega_1 - \omega_2$. These processes described in terms of the electromagnetic waves represent three-photon processes in which the photons $\hbar\omega_1$ and $\hbar\omega_2$ are absorbed and a photon $\hbar\omega = \hbar(\omega_1 + \omega_2)$ is emitted, or the photon $\hbar\omega_1$ is absorbed and photons $\hbar\omega_2$ and $\hbar\omega = \hbar(\omega_1 - \omega_2)$ are emitted, as is obvious from the correspondence principle described in Section 3.5.

The use of the second-order susceptibility for higher-harmonics generation is restricted to those materials without inversion symmetry. As mentioned above, $\chi^{(2)}$ consists of terms containing the products $P_{m\ell'} P_{\ell'\ell} P_{\ell m}$ with two intermediate states

ℓ and ℓ' , and vanishes identically in materials with inversion symmetry where P connects only those states with different parities. In this respect, the *third-order susceptibility* $\chi^{(3)}$ has no such restriction on materials used and has moreover much more varied applications for nonlinear optics because of the participation of four photons. However, it is usually much smaller than $\chi^{(2)}$ because of the smallness of the radiation–matter coupling constant mentioned before, and in order to have large enough $\chi^{(3)}$ it is desirable to choose materials and incident (or generated) photon energy so as to make the energy denominator for the intermediate state as small as possible (a nearly resonant intermediate state). In these higher-order optical responses, the material system, virtually brought into excited states ℓ, ℓ' , etc. in the intermediate steps, returns finally to the initial state m , so that what has really occurred is the conversion of the incident light waves to other light waves. For the purpose of developing a variety of nonlinear optical technologies, a combined use of these *dispersive* optical processes is most useful. See the references in Ref. [3] for more details.

For the spectroscopic study of matter which is the main purpose of this book, however, the *dissipative* optical processes in which the matter is brought to final states different from the initial state are more useful, as we already learned with the one-photon absorption and emission spectra. Coming back to the second-order response, let us consider the optical transition processes in which the second states ℓ' of the material system, to be reached as a result of absorbing two photons ω_1 and ω_2 , or absorbing one photon ω_1 and emitting another photon ω_2 , are *not intermediate* but *final* states n which are different from the initial state m . These two processes, which satisfy the energy conservation $E_n - E_m \equiv \hbar\omega_{nm} = \hbar\omega_1 \pm \hbar\omega_2$ instead of the phase-matching condition, are *two-photon absorption* and *Raman scattering* (*light scattering* is a more general terminology including elastic scattering: $E_n = E_m$), respectively.

Two-photon absorption is sometimes more useful than the ordinary one-photon absorption for the spectroscopic study of electronic excited states of insulators (such as excitons, inclusive of the ionization continuum), since a photon resonating to these intrinsic excitations often has so large a cross-section that it is absorbed within a depth smaller than a wavelength from the surface, being unsuitable for quantitative study of the bulk absorption coefficient. Each of two photons in the transparent region (energy $\hbar\omega_i$ smaller than the lowest exciton energy) can penetrate deep into the crystal without attenuation, while two-photon absorption has so small a cross-section, even when the total energy $\hbar\omega_1 + \hbar\omega_2$ is in the region of intrinsic excitation, that it takes place almost uniformly over the entire thickness of the crystal. There is a similar situation in the Raman scattering process. Here the intermediate states m are electronic excited states (non-resonant to $\hbar\omega_1$) as before, but the final states f (resonant to $\hbar(\omega_1 - \omega_2)$) to be studied are *usually* the excited states of intrinsic

lattice vibrations (host phonons) which are again difficult to measure directly with one-photon spectroscopy because of too large an absorption cross-section.

Historically, Raman scattering preceded two-photon absorption, for the reason described in the historical review by Fröhlich.⁴ For pedagogical reasons, however, we will invert this historical order in the following descriptions.

11.2 Two-photon spectroscopy

Two-photon absorption was theoretically described by Goeppert-Mayer⁵ as early as 1931, but it was not until the appearance of intense laser light that this process was detected in the optical frequency region. Kaiser and Garret⁶ observed two-photon absorption of a 6943 Å ruby laser beam by $\text{CaF}_2 : \text{Eu}^{2+}$ which was transparent for this wavelength, confirming that the intensity of the accompanying luminescence, with wavelength at 4250 Å, is proportional to the square of the intensity of the incident beam. Braunstein⁷ calculated the two-photon absorption coefficient due to excitation of electrons from the valence band to the conduction band, while Loudon⁸ took into account the exciton effect. The first two-photon spectroscopy of excitons was performed by Hopfield *et al.* in KI,⁹ thereby confirming that the two-photon absorption edge is of the forbidden band edge type (second class) in contrast to the one-photon edge which was known to be of the allowed band edge type (first class) (see Section 8.4 (a) and (b)).

The merit of two-photon spectroscopy mentioned towards the end of Section 11.1 was in fact utilized for a quantitative study of exciton linewidths in CuCl ¹⁰ and SnO_2 ,¹¹ including temperature dependence which is difficult to measure directly by one-photon spectroscopy because of too large an absorption coefficient in this region of intrinsic excitation.

Another merit of two-photon absorption is the greater degree of freedom in choosing the energy and momentum of photons and elementary excitations, which enables one to do the spectroscopy in the \mathbf{k} -space as well and hence to directly study the polariton dispersions.¹² This technique is further extended to multiphoton processes as will be mentioned in Section 11.7.

There are further degrees of freedom of photons, their polarization directions. These can be utilized, through their combined use in the multiphoton process, as a tool for studying the symmetries of electronic structures, providing a technique which is much more powerful than that available with the one-photon process. As the simplest example, consider a crystal with inversion symmetry where the dipole matrix element connects between states with different parities. The two-photon transition brings the system to the excited states n (via intermediate states ℓ) with the same parity as the initial state m , which are inaccessible by the one-photon process. This alone provides only complementary knowledge on the electronic structure

of the excited states. However, the dependence of the cross-section of two-photon absorption or Raman scattering on the directions of polarization of the incident and/or scattered light beams as referred to the crystalline axis provides us with much more abundant knowledge on the symmetries of the excited states, especially in isotropic materials where the one-photon absorption spectrum is just isotropic.

To make this point clearer, let us calculate the rate of two-photon absorption by the second-order perturbation theory with $H_1^{(1)}$ given by (3.3.7) as the perturbation. Under the incident beams 1 and 2 of N_λ photons with angular frequency ω_λ and polarization $e_\lambda (\lambda = 1, 2)$, this rate is given, apart from an unimportant factor, by

$$W = N_1 N_2 \delta[E_n - E_m - (\hbar\omega_1 + \hbar\omega_2)] \times |\langle n | e_2 \cdot \mathbf{P} \Lambda(\omega_1) e_1 \cdot \mathbf{P} + e_1 \cdot \mathbf{P} \Lambda(\omega_2) e_2 \cdot \mathbf{P} | m \rangle|^2, \quad (11.2.1)$$

where the $\Lambda(\omega_\lambda)$ s are operators defined by the following summation over all intermediate states $|\ell\rangle$:

$$\Lambda(\omega_\lambda) = \sum_\ell |\ell\rangle (E_\ell - E_m - \hbar\omega_\lambda)^{-1} \langle \ell|, \quad (\lambda = 1, 2). \quad (11.2.2)$$

It is convenient to introduce

$$\Lambda_\pm \equiv \Lambda(\omega_1) \pm \Lambda(\omega_2) \quad (11.2.3)$$

and rewrite the expression within the outer $|\cdots|$ of eq. (11.2.1) as

$$G(e_1, e_2) \equiv \langle n | e_2 \cdot [(\mathbf{P} \Lambda_+ \mathbf{P})_S + (\mathbf{P} \Lambda_- \mathbf{P})_{AS}] \cdot e_1 | m \rangle \quad (11.2.4)$$

where the suffixes S and AS mean the symmetric and antisymmetric parts, respectively, of the tensor (\cdots) , namely, $(\cdots)_{Sij} \equiv [(\cdots)_{ij} + (\cdots)_{ji}]/2$ and $(\cdots)_{ASij} \equiv [(\cdots)_{ij} - (\cdots)_{ji}]/2$.

In order to study the polarization dependence of (11.2.4), we resort to a simple group theoretical argument¹³ (those who are not familiar with group theoretical terminologies (*italicized*) and notations such as Γ can also follow the argument with the help of the geometric or algebraic analogy shown in the parentheses below) and classify the final states n according to its symmetry which is represented by an *irreducible representation* of the *point group* to which the crystal belongs. As a typical system, we consider a crystal of cubic symmetry with point group O_h with a totally symmetric ground state g with one-dimensional representation Γ_1^+ (like the atomic s state). Since any vector \mathbf{P} with three components belongs to the three-dimensional representation Γ_4^- (which behaves like x, y, z), a one-photon transition brings the Γ_1^+ -state through \mathbf{P} to the states with irreducible representation: $\Gamma_1^+ \times \Gamma_4^- = \Gamma_4^-$ (where \times indicates group-theoretic multiplication), while the two-photon transition brings it through $\mathbf{P} \times \mathbf{P}$ to the states with any of four irreducible representations: $\Gamma_1^+ \times \Gamma_4^- \times \Gamma_4^- = \Gamma_1^+ + \Gamma_3^+ + \Gamma_4^+ + \Gamma_5^+$, where Γ_3^+ is a

two-dimensional representation behaving like $x^2 - y^2$, $(2z^2 - x^2 - y^2)/\sqrt{3}$, Γ_4^+ a three-dimensional one behaving like $yz - zy$, $zx - xz$, $xy - yx$, and Γ_5^+ another three-dimensional one behaving like $yz + zy$, $zx + xz$, $xy + yx$. One can then decompose the tensor $[\cdot \cdot \cdot]$ in eq. (11.2.4) into the parts corresponding to the components of the four representations. Denoting the rectangular components of \mathbf{e}_i by ℓ_i, m_i, n_i ($i = 1, 2$), (11.2.4) can be written as follows:

$$\begin{aligned}
 G(\mathbf{e}_1, \mathbf{e}_2) &= (1/3)(\ell_1\ell_2 + m_1m_2 + n_1n_2)(P_x\Lambda_+P_x + P_y\Lambda_+P_y + P_z\Lambda_+P_z) & \Gamma_1^+ \\
 &+ (1/2)(\ell_1\ell_2 - m_1m_2)(P_x\Lambda_+P_x - P_y\Lambda_+P_y) & \\
 &+ (1/2\sqrt{3})(2n_1n_2 - \ell_1\ell_2 - m_1m_2) & \Gamma_3^+ \\
 &\quad \times (1/\sqrt{3})(2P_z\Lambda_+P_z - P_x\Lambda_+P_x - P_y\Lambda_+P_y) & \\
 &+ (1/2)(m_1n_2 - n_1m_2)(P_y\Lambda_-P_z - P_z\Lambda_-P_y) + (\text{cyclic sum.}) & \Gamma_4^+ \\
 &+ (1/2)(m_1n_2 + n_1m_2)(P_y\Lambda_+P_z + P_z\Lambda_+P_y) + (\text{cyclic sum.}). & \Gamma_5^+
 \end{aligned} \tag{11.2.5}$$

The angular dependence of the rate of each transition is therefore given by

$$\begin{aligned}
 \Gamma_1^+ &\rightarrow \Gamma_1^+ : (\ell_1\ell_2 + m_1m_2 + n_1n_2)^2 = (\mathbf{e}_1 \cdot \mathbf{e}_2)^2, \\
 \Gamma_1^+ &\rightarrow \Gamma_3^+ : \ell_1^2\ell_2^2 + m_1^2m_2^2 + n_1^2n_2^2 - (\ell_1\ell_2m_1m_2 + m_1m_2n_1n_2 + n_1n_2\ell_1\ell_2), \\
 \Gamma_1^+ &\rightarrow \Gamma_4^+ : 1 - (\ell_1\ell_2 + m_1m_2 + n_1n_2)^2 = (\mathbf{e}_1 \times \mathbf{e}_2)^2, \\
 \Gamma_1^+ &\rightarrow \Gamma_5^+ : 1 - (\ell_1^2\ell_2^2 + m_1^2m_2^2 + n_1^2n_2^2) \\
 &\quad + 2(\ell_1\ell_2m_1m_2 + m_1m_2n_1n_2 + n_1n_2\ell_1\ell_2).
 \end{aligned} \tag{11.2.6}$$

The angular dependence for all point groups, namely, for crystals of all symmetries, are derived and tabulated in Refs. [14–16].

The first experiment to study the polarization dependence of two-photon absorption spectra was that by Matsuoka on TlCl crystal.¹⁷ He found that the symmetry of the final state is A_{1g} -like from the band edge at 3.56 eV up to 4.2 eV, above which the E_g component starts to mix. It was later pointed out by Fröhlich *et al.*¹⁸ that there are only three linearly independent angular functions among the four expressions of (11.2.6), and that the use of *circularly* (in addition to *linearly*) polarized light beams is necessary for the *unique* decomposition of the final state into individual components. A detailed review on studies of the one- and two-photon spectroscopies and band structures of thallous halides is presented by Kobayashi.¹⁹

In insulators with a dipole-allowed direct bandgap at $\mathbf{k} = \mathbf{0}$ (case (a) of Section 8.4), namely with $\mathbf{p}_{cv}(\mathbf{0}) \equiv (\psi_{c0}^*, (-i\hbar\nabla)\psi_{v0}) \neq 0$, the product of the Bloch functions $\psi_{c0}\psi_{v0}^*$ has Γ_4^- symmetry, and only the excitons with s-like (Γ_1^+) envelope functions F_{ns} and hence with total wave functions $F_{ns}\psi_{c0}\psi_{v0}^*$ are one-photon dipole-allowed ($\Gamma_1^+ \times \Gamma_4^- = \Gamma_4^-$) as already mentioned in Section 8.4, while the p-like (Γ_4^-)

excitons with total symmetry $\Gamma_4^- \times \Gamma_4^- = \Gamma_1^+ + \Gamma_3^+ + \Gamma_4^+ + \Gamma_5^+$ are two-photon allowed with the $n = 1$ exciton line missing. The latter is what was found in two-photon absorptions in alkali halides.^{9,20} In insulators with a dipole-forbidden direct bandgap at $\mathbf{k} = 0$ with $\mathbf{p}_{cv}(\mathbf{0}) = 0$, but with $\nabla_{\mathbf{k}}\mathbf{p}_{cv}(\mathbf{0}) \neq 0$ (case (b) of Section 8.4), only the p-like excitons with total symmetry $\Gamma_4^- \times \Gamma_1^+ = \Gamma_4^-$ are one-photon allowed as was observed in the yellow series excitons of Cu_2O (see Section 8.4). Since the product $\psi_{c0}\psi_{v0}^*$ is an even function in the present case (b), the two-photon-allowed excitons (which should in general be even-parity states) should consist of even-parity envelope functions with angular momentum $l = 0$ (s), 2 (d), etc., in the spherical effective mass approximation, but the more exact symmetry of the exciton states with $n \geq 2$ was found unambiguously to be Γ_5^+ (among other possibilities) by Fröhlich *et al.* with the use of the polarization dependence of two-photon absorption.²¹ To be more elaborate, each of the $n \geq 3$ members consists of doublet lines (as seen in Fig. 11.1 where one- and two-photon spectra are shown for comparison) resulting from the mixing of ns - and nd -states of the yellow series excitons and the n 's-states of the green series (originating from the spin-orbit-split lower valence band) through the electron-hole exchange interaction and other non-spherical perturbation terms. These points were made very clear by Uihlein *et al.* through elaborate theoretical arguments.²²

Two-photon spectroscopy of even-parity states was further extended to the study of *spin-forbidden odd-parity states*²³ by making use of the second-order transition

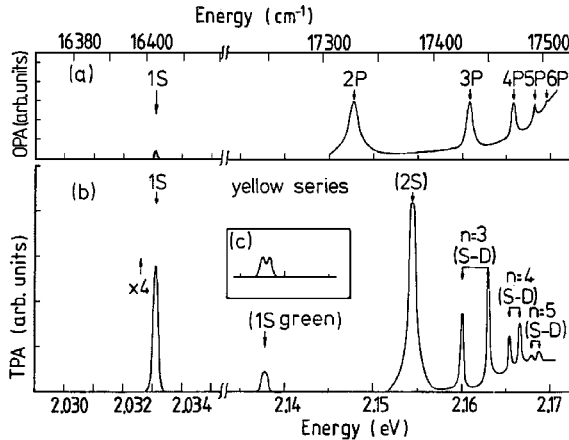


Fig. 11.1 One- and two-photon spectra of Cu_2O . (a) One-photon data at 4.2 K from: J. B. Grun, S. Nikitine, *J. Phys. (Paris)* **23**, 159 (1962); S. Nikitine, J. B. Grun, M. Certier, *Phys. Kondens. Mater.* **1**, 214 (1963). (b) Two-photon data at 4.2 K; the yellow 1s line was measured at 20 K. (c) Splitting of green 1s exciton in a magnetic field of 8.5 T (Voigt configuration). Note the enlarged energy scale for the yellow 1s region. Due to Fröhlich *et al.*²¹

matrix elements in which one of two electric dipoles ED in the electric field ($\propto \mathbf{e} \cdot \mathbf{p}$) is replaced by a magnetic dipole MD in the magnetic field ($\propto i\mathbf{k} \times \mathbf{e} \cdot \mathbf{m}$) (see eqs. (6.7.4, 6.7.5)). In this magnetic–electric dipole two-photon absorption (MED-TPA), one is to replace the expression in the outer $|\cdots\rangle$ of eq. (11.2.1) by

$$\langle f | (MD)_2 \Lambda(\omega_1) (ED)_1 + (MD)_1 \Lambda(\omega_2) (ED)_2 | g \rangle \quad (11.2.7)$$

plus the corresponding terms with (MD) and (ED) being exchanged. The magnetic field of the electromagnetic wave with electric polarization \mathbf{e} and direction of propagation \mathbf{k} is directed along $\mathbf{b} = \mathbf{k} \times \mathbf{e}$, and behaves like an axial vector (vector product) which belongs to the irreducible representation Γ_4^+ . Thus the transition $(MD) \times (ED): (\Gamma_4^+ \times \Gamma_4^-)$ can in general bring the totally symmetric Γ_1^+ ground state to any of the states: $\Gamma_1^- + \Gamma_3^- + \Gamma_4^- + \Gamma_5^-$. If only one laser beam is used for the two-photon transition, the transition $\Gamma_1^+ \rightarrow \Gamma_1^-$ vanishes because \mathbf{b} and \mathbf{e} are perpendicular while the transition $\Gamma_1^+ \rightarrow \Gamma_4^-$ vanishes because $\Lambda_- \equiv \Lambda(\omega_1) - \Lambda(\omega_2) = 0$ (see 11.2.3) and Γ_4^- term in (11.2.5)).

Let us recall the purely triplet exciton state (paraexciton) of alkali halides shown by a dashed line B in Fig. 8.4, which is below the triplet–singlet mixed state (orthoexciton) C by $2/3$ times the (isotropic) exchange energy Δ if $\Delta (>0)$ is small compared to the spin–orbit splitting λ (see eq. (8.5.2)). The state B, which is degenerate in the spherical effective mass approximation used in Section 8.5, in fact consists of Γ_3^- and Γ_5^- in the cubic field, which are split by the *anisotropic* exchange energy ε_{ex} according to Cho *et al.*²⁴ With the use of laser light propagating along the (001) direction with polarizations $\mathbf{e} \parallel (110) (\rightarrow \Gamma_3^-)$ and $\mathbf{e} \parallel (100) (\rightarrow \Gamma_5^-)$, Fröhlich *et al.* obtained directly the splitting ε_{ex} to be 16, 50 and 260 μeV in RbI, NaI and NaBr, respectively, and the corresponding values $\Delta = 33, 17$ and 40 meV from their distance to the transverse 1s exciton C measured by other methods.

M. Itoh and Tanimura²⁵ compared the magnetic field splitting of the paraexciton in NaI measured by the MED-TPA technique mentioned above with the magnetic field shift of the free-exciton luminescence and, on the basis of the agreement between the lower-energy component of the former with the latter, concluded that the optically-excited orthoexcitons are almost exclusively scattered into the paraexciton state during intraband relaxation well before self-trapping. This resolved the long-standing question as to why the singlet self-trapped exciton luminescence is hardly excited when the lowest orthoexcitons are generated by one-photon absorption.

The study of matter with two- and multiphoton processes has been extended to a variety of directions,²⁶ inclusive of the studies of the phase transitions,²⁷ and is now an ordinary and indispensable technique of spectroscopy.

11.3 Light scattering

The elastic and inelastic parts of light scattering are also called *Rayleigh scattering* and *Raman scattering*, respectively. Raman scattering was predicted by Smekal in 1923,²⁸ and experimentally confirmed by Raman in 1928,²⁹ earlier than the prediction as well as the confirmation of two-photon absorption. The lower-energy shift of the scattered light waves from the incident one, $\omega_1 - \omega_2$, is called *Stokes shift*. Scattered light waves with positive and negative Stokes shifts are called *Raman–Stokes* and *Raman–antiStokes lines*, respectively. The dependence of the cross-section of light scattering on the polarization directions of the first (incident) and the second (emitted) photons^{30,31} is similar to that of the two-photon absorption described in the preceding section, except that in (11.2.1, 11.2.2) N_2 and ω_2 are to be replaced by $N_2 + 1$ and $-\omega_2$, respectively. Raman scattering of frequent use in molecules and insulating solids is such that the final states n are vibrationally different from, but electronically the same as, the initial state m while only the intermediate states ℓ are electronically excited states. Namely, the energy difference between the final and initial states, given by $\hbar\omega_{nm} \equiv E_n - E_m = \hbar\omega_1 - \hbar\omega_2$, corresponds to a quantum of molecular or lattice vibrations which is usually much smaller than the electronic excitation energies $\hbar\omega_{\ell m}$ of the intermediate states ℓ . Hereafter within this chapter, we will use the *atomic unit system* with $\hbar = 1$ for the sake of simplifying the quantum-mechanical expressions, and the notations with *capital* letters Ω_1 and Ω_2 for the energies of incident and emitted *photons*, respectively, in order to reserve the *lower case* letter ω for the electronic or vibrational energies of the *material* system.

Being a second-order optical process, Raman scattering has a much smaller cross-section ($\propto W/N_1(N_2 + 1)$ in (11.2.1)) than one-photon absorption. However, the cross-section can be greatly enhanced by making Ω_1 nearly resonant to the intermediate excitations $\omega_{\ell m}$. For such *resonant secondary radiation – resonant Raman scattering* as it is conventionally called – one can neglect the second term in the outer $|\cdots|$ of (11.2.1) against the resonantly-enhanced first term and write down the cross-section as

$$S(\Omega_1, \Omega_2) = \sum_n \left| \sum_{\ell} (P_2)_{n\ell} (P_1)_{\ell m} / (\omega_{\ell m} - \Omega_1) \right|^2 2\pi \delta(\omega_{nm} - \Omega_1 + \Omega_2) \quad (11.3.1)$$

apart from an unimportant factor, where $P_1 \equiv \mathbf{e}_1 \cdot \mathbf{P}$, etc.

Here we come upon a conceptual problem of quantum mechanics – separability of successive elementary processes. One can consider the situation in which the incident photon Ω_1 is absorbed as a *real* (instead of *virtual*) process to raise the system to resonant excited states ℓ ($\omega_{\ell m} - \Omega_1 = 0$) which will sooner or later be de-excited to the states n emitting another photon Ω_2 , again as a real process ($\omega_{n\ell} + \Omega_2 = 0$). Can one distinguish these successive real processes (absorption

and emission) from the Raman scattering which is usually considered to be successive virtual absorption and emission at least in the non-resonant condition? In the resonant situation, are these apparently different descriptions, one as successive two elementary processes of first order and another as a single elementary process of second order, in principle and conceptually distinguishable, or merely different aspects of the same reality? The right answer will turn out to be the latter. Still, one may well ask how the different aspects are logically related. The answer to this intriguing question is not simple, as has been discussed by a number of authors in different situations and by different methods and models.³²⁻⁴⁷

Light scattering itself is a vast field of physics and chemistry. Confining ourselves to the solid state, we can refer to a series of books, *Light Scattering in Solids* edited by Cardona and Güntherodt,⁴⁸ which contain many excellent review articles covering quite a variety of problems. Here in the present book, we will focus our attention on the above-mentioned problem, the effect of resonance, mainly from the theoretical viewpoint.

Let us recapitulate the argument in Section 3.5 where the rate of the one-photon process was derived by the time-integration of the Schrödinger equation (see (3.5.8)) and apply that to the present two-photon process to derive (11.3.1) although it is a well-known general formula for the second-order process. By repeated use of the integrated expression (3.5.8), the amplitude of the final state n reached from the initial state m by first absorbing a photon 1 and then emitting a photon 2 can be written as

$$c_n(t) = C \int_0^t dt_2 \int_0^{t_2} dt_1 \exp[-i(\omega_n + \Omega_2)(t - t_2)] \\ \times \sum_{\ell} (P_2)_{n\ell} \exp[-i\omega_{\ell}(t_2 - t_1)] (P_1)_{\ell m} \exp[-i(\omega_m + \Omega_1)t_1], \quad (11.3.2)$$

where the irrelevant constant C contains factors, other than P_1 and P_2 , of the interactions of matter with radiation fields 1 and 2.

In the second-order perturbation theory (11.3.2), the system is considered to stay in the excited state ℓ from time t_1 until t_2 . More exactly speaking, however, the system is always subject to radiative decay which is of the same origin as we are considering here as the second step but consists of contributions from all possible modes $2'$ of photons and all possible final states n' of the system. Let us take account of this radiative decay phenomenologically by a decay constant γ_{ℓ} appearing as a (negative) imaginary part of energy ω_{ℓ} in units of angular frequency. Since the population of the state ℓ is proportional to its amplitude squared $|\exp[-i(\omega_{\ell} - i\gamma_{\ell})(t_2 - t_1)]|^2 = \exp[-2\gamma_{\ell}(t_2 - t_1)]$, the decay rate is given by $2\gamma_{\ell}$. In order to incorporate this decay into (11.3.2) from first principles of quantum mechanics, one has to consider higher-order terms of the perturbation expansion than the second order (see Ref. [32] for a comprehensive description of the resonance fluorescence

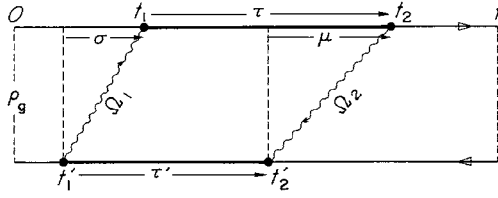


Fig. 11.2 Feynman diagram for resonant second-order optical process. Ω_1 and Ω_2 are energies of incoming and outgoing photons, respectively.⁴⁵

of an atom, and also Ref. [43] for general situations where γ is to be considered as a dynamical variable (operator)).

Taking the statistical average on the initial state m with Boltzmann factor w_m , one can now write the transition probability in the form of the four-fold integral:

$$\begin{aligned} \sum_m w_m |c_n(t)|^2 &= |C|^2 \sum_m \int_0^t dt_2 \int_0^{t_2} dt_1 \int_0^t dt'_2 \int_0^{t'_2} dt'_1 \exp[i(\omega_n + \Omega_2)t_2] \\ &\times \sum_{\ell} (P_2)_{n\ell} \exp[-i(\omega_{\ell} - i\gamma_{\ell})(t_2 - t_1)] (P_1)_{\ell m} \exp[-i(\omega_m + \Omega_1)t_1] \\ &\times w_m \exp[i(\omega_m + \Omega_1)t'_1] \sum_{\ell'} (P_1)_{m\ell'} \exp[i(\omega_{\ell'} + i\gamma_{\ell'})(t'_2 - t'_1)] \\ &\times (P_2)_{\ell'n} \exp[-i(\omega_n + \Omega_2)t'_2], \end{aligned} \quad (11.3.3)$$

as shown schematically in Fig. 11.2. We will study the *asymptotic* behavior of this integral for large enough t , the upper bound of the integration.

Instead of the four time coordinates, it is convenient to make use of a *mean* time

$$s \equiv (t_1 + t_2 + t'_1 + t'_2)/4 \quad (11.3.4)$$

and four *relative* times

$$\tau \equiv t_2 - t_1, \quad \tau' \equiv t'_2 - t'_1, \quad \sigma \equiv t_1 - t'_1, \quad \mu \equiv t_2 - t'_2 (= \sigma + \tau - \tau') \quad (11.3.5)$$

of which three are independent. The four-fold integration can now be rewritten as

$$\int ds \int d\tau \int d\tau' \int d\sigma \quad (11.3.6)$$

since the Jacobian $\partial(s, \tau, \tau', \sigma)/\partial(t_1, t_2, t'_1, t'_2)$ for this transformation of variables is unity. However, the region to be integrated on the new time variables is somewhat complicated.

The product of the four exponentials in (11.3.3) is now written as

$$\begin{aligned} &\exp[i(\omega_{nm} - \Omega_1 + \Omega_2)\sigma] \exp[-i(\omega_{\ell n} - i\gamma_{\ell} - \Omega_2)\tau] \\ &\times \exp[i(\omega_{\ell'n} + i\gamma_{\ell'} - \Omega_2)\tau']. \end{aligned} \quad (11.3.7)$$

Let us first note that the second and third exponentials decay with increasing τ and τ' due to the finite decay constants γ_{ℓ} . The integration on τ and τ' converges well

before the right hand ends of the trapezoidal diagram of Fig. 11.2 reaches the upper bound t , provided that t is taken to be large enough (much larger than any of γ_ℓ^{-1}). Let us put s (the center of the trapezoid) to be remote enough ($\gg \gamma_\ell^{-1}$) from the lower (0) and upper (t) bounds. Then the integration on σ gives

$$\int_{-\infty}^{+\infty} d\sigma \exp[i(\omega_{nm} - \Omega_1 + \Omega_2)\sigma] = 2\pi \delta(\omega_{nm} - \Omega_1 + \Omega_2). \quad (11.3.8)$$

The remaining integration on s gives approximately t (the upper bound) if the latter is large enough, since the integrand is independent of s . In this way, the transition rate, which is the transition probability (11.3.6) divided by t , can be calculated, through some rearrangements, as

$$\begin{aligned} \text{Lim}_{t \rightarrow \infty} \left[\sum_m w_m |c_n(t)|^2 / t \right] &= |C|^2 \int_0^\infty d\tau \int_0^\infty d\tau' \int_{-\infty}^\infty d\sigma \\ &\times \exp[-i\Omega_1\sigma + i\Omega_2\mu] \sum_n \sum_m \sum_{\ell'} (P_1)_{m\ell'} \exp[i(\omega_{\ell'} + i\gamma_{\ell'})\tau'] (P_2)_{\ell'n} \\ &\times \exp[i\omega_n\mu] \sum_{\ell} (P_2)_{n\ell} \exp[-i(\omega_{\ell} - i\gamma_{\ell})\tau] (P_1)_{\ell m} \exp[-i\omega_m\sigma] w_m \\ &= |C|^2 \sum_n \sum_m w_m \left| \sum_{\ell} (P_2)_{n\ell} (P_1)_{\ell m} / (\omega_{\ell m} + i\gamma_{\ell} - \Omega_1) \right|^2 \\ &\times 2\pi \delta(\omega_{nm} - \Omega_1 + \Omega_2) \equiv |C|^2 S_\gamma(\Omega_1, \Omega_2). \end{aligned} \quad (11.3.9)$$

$S_\gamma(\omega_1, \omega_2)$ is nothing other than the expression (11.3.1) for the second-order process except for the inclusion of the decay term γ_ℓ as the imaginary part of the energy (in angular frequency) ω_ℓ of the excited state. The inclusion is essential for the convergence of the rate of the *resonant* second-order process. Otherwise the transition probability (11.3.3) would increase as t^2 instead of t due to the τ - and τ' -integrations provided there are excited states ℓ resonant to the incident light: $\omega_{\ell m} - \omega_1 (= \omega_{\ell n} - \omega_2) = 0$ (see (11.3.7)).

If one defines the Hamiltonians for the ground- and excited-state subspaces by

$$H_g \equiv \sum_m |m\rangle \omega_m \langle m| \quad (n \text{ is included among the } ms)$$

and

$$H_e \equiv \sum_{\ell} |\ell\rangle \omega_{\ell} \langle \ell|, \quad (11.3.10)$$

respectively, the density matrix for the ground states by

$$\rho_g \equiv \sum_m |m\rangle w_m \langle m| \quad (11.3.11)$$

and the operator for the radiative energy width in the excited-state subspace by

$$\gamma \equiv \sum_{\ell} |\ell\rangle \gamma_{\ell} \langle \ell|, \quad (11.3.12)$$

one can write the relevant part of (11.3.9) in a more compact form:

$$\begin{aligned}
 S_\gamma(\Omega_1, \Omega_2) = & \int_0^\infty d\tau \int_0^\infty d\tau' \int_{-\infty}^\infty d\sigma \\
 & \times \text{Tr}_g \{ P_1 \exp[i(H_e + i\gamma)\tau'] P_2 \exp[i(H_g + \Omega_2)\mu] \\
 & \times P_2 \exp[-i(H_e - i\gamma)\tau] P_1 \exp[-i(H_g + \Omega_1)\sigma] \rho_g \} \quad (11.3.13)
 \end{aligned}$$

The *trace* operation Tr_g indicates to take summation of all diagonal elements of the propagator in $\{\dots\}$ within the ground-state subspace: $\sum_m \{\dots\}_{mm}$.

The propagator \dots to be read from right to left corresponds to the trapezoid in Fig. 11.2 to be traced clockwise. Starting from the lower left corner of the trapezoid at time t_1' , the system propagates in the ground state (H_g) interacting through P_1 with the incident photon Ω_1 until time t_1 when it is excited, absorbing the photon, then it propagates with energy H_e being always subject to the radiative decay through the imaginary part γ until time t_2 when it is de-excited, emitting a photon Ω_2 through the interaction P_2 , and then repeats the same processes in opposite order and in the opposite time direction corresponding to the conjugate complex of the wave function.

Let us first consider a two-level atom with a non-degenerate ground state ($\omega_m = \omega_g$) and excited state ($\omega_\ell = \omega_e = \omega_g + \varepsilon$), for which (11.3.13) gives

$$S_\gamma(\Omega_1, \Omega_2) = 2\pi |(P_1)_{eg}|^2 |(P_2)_{ge}|^2 \delta(\Omega_2 - \Omega_1) [(\Omega_2 - \varepsilon)^2 + \gamma^2]^{-1}, \quad (11.3.14)$$

as was described in Ref. [32]. For monochromatic incident light with intensity $I(\Omega_1) = I_0 \delta(\Omega_1 - \Omega_0)$, we have the emitted light with intensity

$$\begin{aligned}
 I(\Omega_2) & \propto \int d\Omega_1 S_\gamma(\Omega_1, \Omega_2) I(\Omega_1) \\
 & \propto I_0 \delta(\Omega_2 - \Omega_0) [(\Omega_0 - \varepsilon)^2 + \gamma^2]^{-1}. \quad (11.3.15)
 \end{aligned}$$

Namely, it is elastic scattering (also called Rayleigh scattering) with cross-section proportional to the absorption spectrum of the Lorentzian shape. The atom has only been virtually (not really) excited during this scattering process, while the photon has only been scattered as is obvious from $\Omega_2 = \Omega_0$. For pulsed and hence white (due to the uncertainty principle) incident light with $I(\Omega_1) = I$ (constant), we have the emitted light with intensity

$$I(\Omega_2) \propto I [(\Omega_2 - \varepsilon)^2 + \gamma^2]^{-1}, \quad (11.3.16)$$

indicating that the atom has once absorbed an incident photon Ω_1 as a real process, and re-emits another photon Ω_2 after it has lost the memory of which photon energy it has absorbed ($\Omega_2 \neq \Omega_1$ although both of them are within $\sim \pm\gamma$ from ε). The interpretation of what has taken place in these opposite limits of incident light is most clear, one being scattering and another being absorption followed by emission.

Although the interpretation becomes obscured in the intermediate situations of incidence, it is obvious that the two interpretations are different aspects of the same reality. It is left for the interested reader to confirm this statement with the use of the time-dependent description (11.3.13) applied to a (temporally as well as spectrally) general form of incidence.

In the general case where both H_g and H_e have internal structures (consist of many states), let us note that (11.3.13) can be rewritten, by changing the variable σ to μ with the use of (11.3.5), as

$$S_\gamma(\Omega_1, \Omega_2) = \int_{-\infty}^{\infty} d\mu \exp[i(\Omega_2 - \Omega_1)\mu] \langle \langle \alpha_{\Omega_1}^\dagger \alpha_{\Omega_1}(\mu) \rangle \rangle \quad (11.3.17)$$

where α_{Ω_1} is an operator defined by

$$\alpha_{\Omega_1} \equiv \int_0^\infty i d\tau \exp(i\Omega_1 \tau) \{ P_2 \exp[-i(H_e - \gamma)\tau] P_1 \exp[iH_g \tau] \}, \quad (11.3.18)$$

$\alpha_{\Omega_1}(\mu) = \exp(iH_g \mu) \alpha_{\Omega_1} \exp(-iH_g \mu)$ is its Heisenberg representation at time μ , and $\langle \langle \cdots \rangle \rangle \equiv \text{Tr}_g \{ \cdots \rho_g \}$. Note that

$$\langle \langle \alpha_{\Omega_1} \rangle \rangle = \mathbf{e}_1 \cdot \overline{\overline{\chi}}(\Omega_1) \cdot \mathbf{e}_2 \quad (11.3.19)$$

where $\overline{\overline{\chi}}(\Omega_1)$ is the susceptibility tensor given by

$$\chi(\Omega_1)_{ij} = \sum_m \sum_\ell w_m (P_1)_{m\ell} (\omega_{\ell m} - i\gamma_\ell - \Omega_1)^{-1} (P_j)_{\ell m} \quad (11.3.20)$$

as is obtained from the tensor components and matrix elements of (11.3.18) by performing the time integration. It is in fact the first term (resonant part) of (6.1.31) with infinitesimal γ replaced by finite decay constant γ_ℓ and Fourier component \mathbf{P}_k by \mathbf{P}_0 . The operator α_{Ω_1} defined by (11.3.18) will be called the *polarizability* operator to distinguish it from susceptibility which is its expected value.

Equation (11.3.17) can be decomposed as

$$\begin{aligned} S_\gamma(\Omega_1, \Omega_2) &= 2\pi \delta(\Omega_2 - \Omega_1) \langle \langle \alpha_{\Omega_1}^\dagger \rangle \rangle \langle \langle \alpha_{\Omega_1} \rangle \rangle \\ &+ \int_{-\infty}^{\infty} d\mu \exp[i(\Omega_2 - \Omega_1)\mu] \langle \langle \{ \alpha_{\Omega_1}^\dagger - \langle \langle \alpha_{\Omega_1}^\dagger \rangle \rangle \} \{ \alpha_{\Omega_1}(\mu) - \langle \langle \alpha_{\Omega_1} \rangle \rangle \} \rangle \rangle \end{aligned} \quad (11.3.21)$$

under the assumption of the ergodicity within the subspace g :

$$\text{Lim}_{|\mu| \rightarrow \infty} \langle \langle \alpha_{\Omega_1}^\dagger \alpha_{\Omega_1}(\mu) \rangle \rangle = \langle \langle \alpha_{\Omega_1}^\dagger \rangle \rangle \langle \langle \alpha_{\Omega_1} \rangle \rangle \quad (11.3.22)$$

which is usually expected to be valid for large systems.

Equation (11.3.21) means that the average and the fluctuation therefrom of polarizability are responsible for the elastic and inelastic parts, respectively, of

light scattering, namely, the Rayleigh and Raman scattering. This is in accordance with the semi-classical picture for light scattering which was already considered in non-resonant situations.⁴⁹ Namely, the incident light Ω_1 induces oscillating electronic polarization with proportionality constant α_{Ω_1} which in turn emits the secondary radiation. If α_{Ω_1} itself, as a dynamical variable, has vibrating components (for instance, the electronic polarizability of a diatomic molecule depends on the interatomic distance which is vibrating), the frequency Ω_2 of the thereby scattered light will be shifted from Ω_1 by this vibrational frequency. Integrating (11.3.17 and 11.3.21) on Ω_2 , we find that the fractional intensity of Rayleigh scattering is given by

$$R_0 = \langle\langle \alpha_{\Omega_1}^\dagger \rangle\rangle \langle\langle \alpha_{\Omega_1} \rangle\rangle / \langle\langle \alpha_{\Omega_1}^\dagger \alpha_{\Omega_1}(\mu) \rangle\rangle. \quad (11.3.23)$$

The above decomposition of the four-time correlation function (11.3.13) into the two-time correlation function (11.3.17) of polarizability, as schematically shown in the upper half of Fig. 11.3, is one aspect of the resonant second-order optical process. There is an alternative decomposition into two-time correlation functions for absorption and emission processes, as shown in the lower half of the figure. The complementarity of these *dual* aspects of the resonant second-order process was discussed in Ref. [45]. In the following two sections, we will study these two aspects of the secondary radiation using a localized electron interacting with lattice vibrations as a model system.

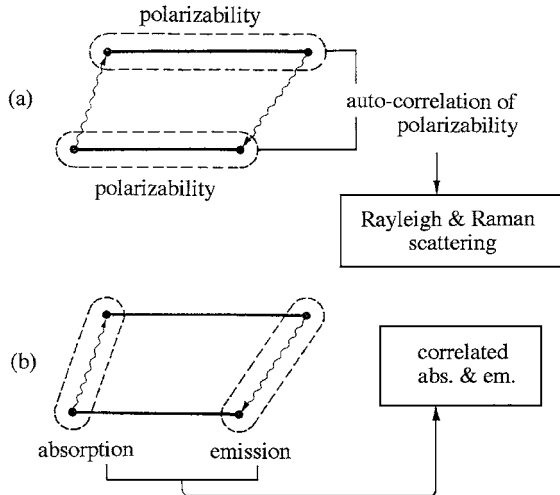


Fig. 11.3 Two complementary ways of decomposing the second-order optical process. (a) Light scattering in terms of auto-correlation of polarizability. Correlated absorption and emission processes.⁴⁵

11.4 Resonant light scattering by a localized electron

As typical systems for resonant light scattering, we will consider (i) a two-level localized electron in this and the following section, and (ii) an exciton in Section 11.6, both being *microscopic* entities interacting with lattice vibrations as their *macroscopic* background. Stochastic models for the latter as an external source (like a heat bath) of fluctuation of various *specified* characteristics provide sometimes illuminating answers to the questions mentioned below (11.3.1). However, we will take here the dynamical description of both the microscopic and macroscopic systems on an equal footing since we want to take into account a finite reaction of the latter – lattice relaxation – to the optical excitation of the former.

We denote the ground and excited states of a localized electron by $|g\rangle$ and $|e\rangle$, and their adiabatic potentials by $W_g(Q)$ and $W_e(Q)$, respectively, where Q represents the displacements of all atoms of the lattice. Adding the kinetic energy of the lattice, K_L , to W_g , we obtain the Hamiltonian $H_L = K_L + W_g$ for the lattice vibrations when the electron is in the ground state. Let us decompose the adiabatic excitation energy $\Delta W(Q) \equiv W_e(Q) - W_g(Q)$ into the electronic part ε_a and deviation $\delta V(Q)$ therefrom. How to decompose it will be mentioned later. The deviation $\delta V(Q)$ will be called the interaction of the electron with the lattice vibrations. One can then write the total Hamiltonian of the system as

$$H_g = |g\rangle H_L \langle g|, \quad H_e = |e\rangle (H_L + E_a + \delta V) \langle e|. \quad (11.4.1)$$

While $W_g(Q)$ and $W_e(Q)$ depend on the displacements of atoms of the entire lattice, the electron-originated difference $\Delta W(Q)$ and hence $\delta V(Q)$ depend only on those of the atoms within the spatial extension of the electron. If, moreover, one assumes $W_g(Q)$ to be harmonic and $\delta V(Q)$ to be linear in the atomic displacements Q , the system reduces to that described in Chapter 4, a harmonic lattice with linear electron–lattice interaction. However, we do not specify the model here beyond the Hamiltonian (11.4.1) with macroscopically extended H_L and microscopically localized $\delta V(Q)$.

With the use of (11.4.1), (11.3.13) can be written as

$$\begin{aligned} S_\gamma(\Omega_1, \Omega_2) = & |(P_1)_{eg}|^2 |(P_2)_{eg}|^2 \int_0^\infty d\tau \int_0^\infty d\tau' \int_{-\infty}^\infty d\sigma \\ & \times \langle \langle \exp[i(H_L + E_a + \delta V + i\gamma)\tau'] \exp[i(H_L + \Omega_2)\mu] \\ & \times \exp[-i(H_L + E_a + \delta V - i\gamma)\tau] \exp[-i(H_L + \Omega_1)\sigma] \rangle \rangle, \end{aligned} \quad (11.4.2)$$

$$\text{where } \langle \langle \cdots \rangle \rangle \equiv \text{Tr}\{\cdots \rho_g\}, \quad \rho_g = \exp(-\beta H_L) / \text{Tr}[\exp(-\beta H_L)].$$

$$(11.4.3)$$

Let us now define

$$f(\sigma) \equiv \langle \langle \exp[i(H_L \sigma)] \exp[-i(H_L + E_a + \delta V)\sigma] \rangle \rangle = \exp(-i E_a \sigma) \langle \langle S_\sigma \rangle \rangle \quad (11.4.4)$$

$$\text{where } S_\sigma \equiv \exp[i H_L \sigma] \exp[-i(H_L + \delta V)\sigma]. \quad (11.4.5)$$

The Fourier transform of $f(\sigma)$ gives the normalized absorption spectrum (without the γ -effect) of this localized electron:

$$\begin{aligned} F_a(\Omega_1) &= (2\pi)^{-1} \int_{-\infty}^{\infty} \exp(i \Omega_1 \sigma) f(\sigma) d\sigma \\ &= (2\pi)^{-1} \int_{-\infty}^{\infty} \exp[-i(\Omega_1 - E_a)\sigma] \langle \langle S_\sigma \rangle \rangle d\sigma. \end{aligned} \quad (11.4.6)$$

This can be seen by denoting the eigenvalues and eigenstates of H_L by ω_m and $|m\rangle$, respectively, those of $H_L + E_a + \delta V$ by ω_ℓ and $|\ell\rangle$, and explicitly writing down (11.4.6) as

$$F_a(\Omega_1) = \sum_m \sum_\ell w_m |\langle m|\ell \rangle|^2 \delta(\omega_\ell - \omega_m - \Omega_1). \quad (11.4.6a)$$

Since S_σ satisfies the differential equation

$$(d/d\sigma) S_\sigma = -i \exp[i H_L \sigma] \delta V \exp[-i(H_L + \delta V)\sigma] = -i \delta V(\sigma) S_\sigma, \quad (11.4.7)$$

where

$$\delta V(\sigma) \equiv \exp(i H_L \sigma) \delta V \exp(-i H_L \sigma), \quad (11.4.8)$$

it can be written, with the repeated integration of (11.4.7), as an infinite series:

$$\begin{aligned} S_\sigma &= 1 - i \int_0^\sigma d\sigma_1 \delta V(\sigma_1) S_{\sigma_1} \\ &= 1 - i \int_0^\sigma d\sigma_1 \delta V(\sigma_1) - \int_0^\sigma d\sigma_1 \int_0^{\sigma_1} d\sigma_2 \delta V(\sigma_1) \delta V(\sigma_2) + \dots \end{aligned} \quad (11.4.9)$$

By choosing E_a , so far unspecified, in such a way that the average of fluctuation δV on the ensemble (and hence on time since H_L commutes with ρ_g) vanishes:

$$\langle \langle \delta V(\sigma) \rangle \rangle = 0, \quad (11.4.10)$$

the average of (11.4.9) can be expanded as

$$\begin{aligned} \langle \langle S_\sigma \rangle \rangle &= 1 - \int_0^\sigma d\tau (\sigma - \tau) \langle \langle \delta V(\tau) \delta V \rangle \rangle + (\text{higher-order terms}), \quad (11.4.11) \\ &= \exp \left[- \int_0^\sigma d\tau (\sigma - \tau) \langle \langle \delta V(\tau) \delta V \rangle \rangle + (\text{higher-order cumulants}) \right], \end{aligned} \quad (11.4.11a)$$

namely, in terms of the second- and higher-order correlation functions or cumulants⁵⁰ of δV . In the case of a harmonic lattice with linear electron–lattice interaction as described in Chapter 4, the higher-order cumulants in (11.4.11a) exactly vanish as can be readily confirmed below.

Let us decompose the correlation function into the Fourier components as

$$\langle\langle\delta V(\tau)\delta V\rangle\rangle = \int_{-\infty}^{\infty} J(\omega) \exp(-i\omega\tau) d\omega. \quad (11.4.12)$$

Putting this into the integral in (11.4.1) and the thus-obtained $\langle\langle S_\sigma \rangle\rangle$ into (11.4.5), we can write $f(\sigma)$ as

$$\begin{aligned} f(\sigma) = & \exp[-i\sigma\{E_a - \int_{-\infty}^{\infty} d\omega J(\omega)\omega^{-1}\} \\ & - \int_{-\infty}^{\infty} d\omega J(\omega)\omega^{-2} + \int_{-\infty}^{\infty} d\omega J(\omega)\omega^{-2} \exp(-i\omega\sigma)]. \end{aligned} \quad (11.4.13)$$

This can be identified term by term with eq. (4.5.10a), which is an exact expression for the harmonic lattice with linear electron–lattice interaction, by noting the following correspondence:

$$\begin{aligned} J(\omega)\omega^{-2} & \Leftrightarrow s(\omega)\{n_\beta(\omega) + 1\}, \quad (n_\beta(\omega) = [\exp(\beta\omega) - 1]^{-1}) \\ \int_{-\infty}^{\infty} d\omega J(\omega)\omega^{-1} & \Leftrightarrow \int_0^{\infty} s(\omega)\omega d\omega = E_R, \\ \{\dots\} \text{ in (11.4.13)} & \Leftrightarrow E_0 \end{aligned} \quad (11.4.14)$$

The expression $\langle\langle \dots \rangle\rangle$ in (11.4.2), after separation of c-number factor, can now be rewritten in terms of this S_σ , defined by (11.4.6), and its Heisenberg representation $S_\sigma(\mu)$ defined by (11.4.8), as

$$\begin{aligned} & \langle\langle \exp[i(H_L + \delta V)\tau'] \exp[iH_L\mu] \exp[-i(H_L + \delta V)\tau] \exp[-iH_L\sigma] \rangle\rangle \\ & = \langle\langle S_{-\tau'} S_{-\tau}^\dagger(\mu) \rangle\rangle. \end{aligned} \quad (11.4.15)$$

Inserting (11.4.9) and making use of the cumulant expansion technique as was done in (11.4.11a), one obtains, following Hizhnyakov and Tehver³⁴ who presented a pioneer work on this problem,

$$\langle\langle S_{-\tau'} S_{-\tau}^\dagger(\mu) \rangle\rangle = \langle\langle S_{-\tau'} \rangle\rangle \langle\langle S_{-\tau}^\dagger \rangle\rangle \exp[K(\tau, \tau', \mu) + (\text{higher-order cumulants})], \quad (11.4.16)$$

where

$$K(\tau, \tau', \mu) \equiv \int_0^{-\tau'} ds' \int_0^{-\tau} ds \langle\langle \delta V(s') \delta V(s + \mu) \rangle\rangle \quad (11.4.17)$$

$$\begin{aligned} & = \int_{-\infty}^{\infty} d\omega J(\omega)\omega^{-2} \exp(i\omega\mu) [1 - \exp(-i\omega\tau)] \\ & \quad \times [1 - \exp(i\omega\tau')]. \end{aligned} \quad (11.4.18)$$

$$J(\omega) \equiv (2\pi)^{-1} \int_{-\infty}^{\infty} \exp(-i\omega t) \langle \langle \delta V \delta V(t) \rangle \rangle dt. \quad (11.4.19)$$

It is noted in (11.4.16) again that the higher-order cumulants vanish for a harmonic lattice with linear electron–lattice interaction.

Noting that $\langle \langle S_{-\tau} \rangle \rangle = \exp(-iE_a\tau)f(-\tau)$, one can write (11.4.2) as

$$\begin{aligned} S_\gamma(\Omega_1, \Omega_2) &= |(P_1)_{\text{eg}}|^2 |(P_2)_{\text{eg}}|^2 \int_0^\infty d\tau \int_0^\infty d\tau' \int_{-\infty}^\infty d\mu \\ &\quad \times \exp[i(\Omega_1 - \gamma)\tau + (-i\Omega_1 - \gamma)\tau' + i(\Omega_2 - \Omega_1)\mu] \\ &\quad \times f^*(-\tau)f(-\tau') \sum_{n=0}^\infty [K(\tau, \tau', \mu)]^n / n! \\ &= \sum_n S_n(\Omega_1, \Omega_2). \end{aligned} \quad (11.4.20)$$

Let us note that the susceptibility component relevant to the light scattering is given by

$$\langle \langle \alpha_{\Omega_1} \rangle \rangle / (P_2)_{\text{ge}}(P_1)_{\text{eg}} \equiv \Phi_\gamma(\Omega_1) = \int_{-\infty}^\infty d\Omega F(\Omega) [\Omega - \Omega_1 - i\gamma]^{-1} \quad (11.4.21)$$

as can be seen from (11.3.18) with the use of (11.4.1, 11.4.4) and the inverse transform of (11.4.6). Then n th order Raman term of (11.4.20) can be rewritten as

$$\begin{aligned} S_n(\Omega_1, \Omega_2) &= |(P_2)_{\text{ge}}(P_1)_{\text{eg}}|^2 \int_{-\infty}^\infty J(\omega_1) d\omega_1 \cdots \int_{-\infty}^\infty J(\omega_n) d\omega_n \\ &\quad \times \delta(\Omega_1 - \Omega_2 - \omega_1 - \cdots - \omega_n) (n!)^{-1} \left| \left\{ \prod_{v=1}^n \partial(\omega_v) \right\} \Phi_\gamma(\Omega_1) \right|^2 \end{aligned} \quad (11.4.22)$$

where $\partial(\omega_v)$ is the difference operator defined by

$$\partial(\omega)\Phi_\gamma(\Omega) \equiv [\Phi_\gamma(\Omega) - \Phi_\gamma(\Omega - \omega)]/\omega. \quad (11.4.23)$$

The zeroth-order term represents the Rayleigh scattering. While an expression equivalent to eq. (11.4.22) was given in Ref. [51], we give here the present form in terms of the difference operator for later use.

Let us now study the two limiting situations of the temporal behavior of the correlation function $\langle \langle \delta V \delta V(t) \rangle \rangle$ characterized by the two parameters, the amplitude D and the correlation time τ_c , defined as:

$$D^2 \equiv \langle \langle (\delta V)^2 \rangle \rangle, \quad (11.4.24)$$

$$D^2 \tau_c \equiv \int_0^\infty \langle \langle \delta V \delta V(t) \rangle \rangle dt \equiv \Gamma = \pi J(0). \quad (11.4.25)$$

The two limiting situations are characterized as follows:

$$D\tau_c \gg 1 \text{ (slow modulation, or strong coupling),} \quad (11.4.26)$$

$$D\tau_c \ll 1 \text{ (rapid modulation, or weak coupling, or motional narrowing).} \quad (11.4.27)$$

Let us note that the integral in (11.4.11a) can be approximated as

$$\int_0^\sigma d\tau (\sigma - \tau) \langle \langle \delta V \delta V(\tau) \rangle \rangle \sim D^2 \sigma^2 / 2 \text{ or } \sim \Gamma \sigma$$

according as $\sigma \ll \tau_c$ or $\sigma \gg \tau_c$. The lineshape of the absorption spectrum, (11.4.6), is governed by the behavior of the integral in the σ -region where it is of the order of unity (provided the higher-order cumulants are negligible), and is therefore given by

$$F_a(\Omega_1) = (2\pi D^2)^{-1/2} \exp[-(\Omega_1 - E_a)^2 / 2D^2] \quad (\text{slow modulation}), \quad (11.4.28)$$

$$= \pi^{-1} \Gamma [(\Omega_1 - E_a)^2 + \Gamma^2]^{-1} \quad (\text{rapid modulation}). \quad (11.4.29)$$

The successive Raman terms in the rapid-modulation case are calculated as follows. Putting (11.4.29) into (11.4.21) and (11.4.23), we obtain

$$\Phi_\gamma(\Omega) [\varepsilon - \Omega - i(\Gamma + \gamma)]^{-1}, \quad (11.4.30)$$

$$\partial(\omega_v) \Phi_\gamma(\Omega) \equiv \Phi_\gamma(\Omega) \Phi_\gamma(\Omega - \omega_v). \quad (11.4.31)$$

With the repeated use of eq. (11.4.31), we obtain

$$\left\{ \prod_{v=1}^n \partial(\omega_v) \right\} \Phi_\gamma(\Omega) = \Phi_\gamma(\Omega) \Phi_\gamma(\Omega - \omega_1 - \omega_2 - \dots - \omega_v) \\ \times \sum_{n!} \Phi_\gamma(\Omega - \omega_{v_1}) \Phi_\gamma(\Omega - \omega_{v_1} - \omega_{v_2}) \dots \Phi_\gamma(\Omega - \omega_{v_1} \dots - \omega_{v_{n-1}}), \quad (11.4.32)$$

where the summation runs over all the possible permutations (${}_n P_{n-1} = n!$) of $(n-1)$ numbers: v_1, v_2, \dots, v_{n-1} out of $(1, 2, \dots, n)$. In calculating the contribution of the square of each summand in eq. (11.4.29) to the integral in eq. (11.4.19), we note that $J(\omega_v)$ can be taken to be constant $\sim J(0) = \Gamma/\pi$ within the region $\omega_v = O(\Gamma + \gamma)$ where $|\Phi_\gamma|^2$ has a significant value, so long as

$$\Gamma \tau_c (= D^2 \tau_c^2) \ll 1 \text{ and } \gamma \tau_c \ll 1. \quad (11.4.33)$$

The contribution of each cross term vanishes in the same approximation. In this way, we have

$$S_n(\Omega_1, \Omega_2) = 2 |(P_2)_{\text{ge}}(P_1)_{\text{eg}}|^2 \Gamma^n / (\Gamma + \gamma)^{n-1} \\ \times [(\Omega_1 - E_a)^2 + (\Gamma + \gamma)^2]^{-1} [(\Omega_2 - E_a)^2 + (\Gamma + \gamma)^2]^{-1}, \quad (11.4.34)$$

$$S_0(\Omega_1, \Omega_2) = 2 |(P_2)_{\text{ge}}(P_1)_{\text{eg}}|^2 [(\Omega_1 - E_a)^2 + (\Gamma + \gamma)^2]^{-1} \pi \delta(\Omega_1 - \Omega_2),$$

with fractional intensities given by

$$R_n = \gamma \Gamma^n / (\Gamma + \gamma)^{n+1} \quad (n \geq 0). \quad (11.4.35)$$

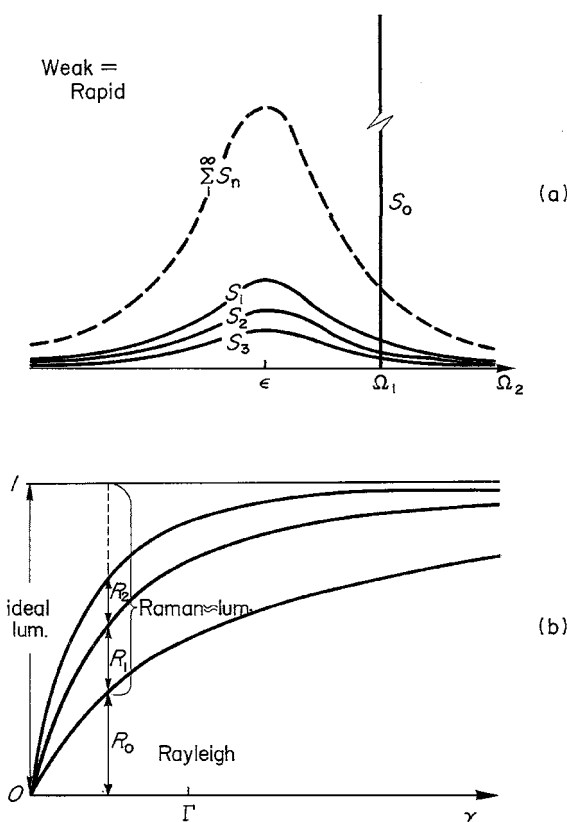


Fig. 11.4 (a) Resonance Raman spectra of successive orders, and (b) their fractional intensities, for the weak coupling case (lum.: luminescence).⁴⁵

The result (11.4.31 to 11.4.32) is a simplified version of those which were obtained by the resolvent method applied to a particular model of quadratic electron–lattice interaction with the phonon structures being retained in the width function $\Gamma(\omega)$.⁴⁴

The spectral behaviors and the fractional intensities of successive Raman scatterings given by (11.4.31) and (11.4.32) are shown in Fig. 11.4(a) and (b), respectively. The characteristic feature of light scattering in the rapid-modulation case is that already the first- (and hence, also higher-) order Raman scattering behaves essentially like the luminescence. The intensity ratio of the sum of Raman scattering of all orders, which is to be interpreted as luminescence, to the Rayleigh scattering is given by $\Gamma : \gamma$, in accordance with the result obtained by other methods such as stochastic models.^{35,42,46}

In contrast, Raman spectra of successive orders in the case of slow modulation gradually change their nature, through hot luminescence, towards the ordinary

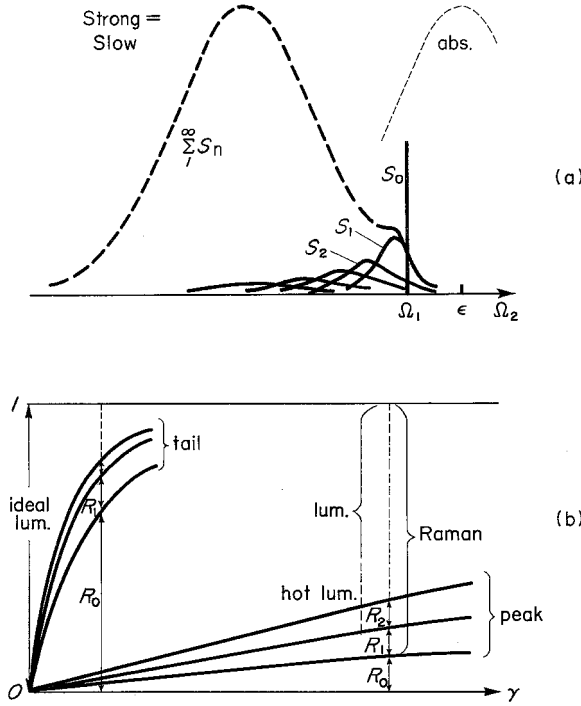


Fig. 11.5 (a) Resonance Raman spectra of successive orders, and (b) their fractional intensities, for the strong coupling case (lum.: luminescence; abs.: absorption).⁴⁵

luminescence, as shown schematically in Fig. 11.5(a) and (b). In fact, the difference operator defined by (11.4.23) can be approximated by a differential operator since the width D of the absorption band (11.4.28) is much larger than the typical phonon energy ω_v . We have therefore

$$\begin{aligned}
 S_n(\Omega_1, \Omega_2) &\simeq |(P_2)_{\text{ge}}(P_1)_{\text{eg}}|^2 |\Phi_\gamma^{(n)}(\Omega_1)|^2 / n! \\
 &\times \int_{-\infty}^{\infty} J(\omega_1) d\omega_1 \dots \int_{-\infty}^{\infty} J(\omega_n) d\omega_n 2\pi \delta(\Omega_1 - \Omega_2 - \omega_1 - \dots - \omega_n).
 \end{aligned}
 \quad (11.4.36)$$

It should be noted that the spectral shape of the n th-order term as a function of energy shift $\Omega_1 - \Omega_2$ is given by the n -fold convolution of the fluctuation spectrum $J(\omega)$. Because of the relation $J(\omega)/J(-\omega) = \exp(\beta\omega)$, generally valid for a quantum-statistical system (see also the first line of eq. (11.4.14)), $J(\omega)$ and the Stokes-shift $\Omega_1 - \Omega_2$ in the Raman spectra are of positive predominance (the Stokes lines ($\Omega_1 - \Omega_2 > 0$) are always stronger than the corresponding anti-Stokes lines). The Ω_1 dependence appears only through the relative intensity of

successive terms:

$$R_n \propto |D^n \Phi_\gamma^{(n)}(\Omega_1)|^2 / n!. \quad (11.4.37)$$

The ratio of successive terms, R_{n+1}/R_n is of the order of unity for the resonance case ($|\Omega_1 - E_a| < D$) while it is of the order of $D/|\Omega_1 - E_a|$ for off-resonance, provided $\gamma < D$.

It is to be noted, however, that the approximation (11.4.36) based on the use of differential operator for (11.4.23) no longer holds when n times the phonon energy becomes comparable to D since $\Phi_\gamma(\Omega_1)$ cannot be taken to be a slowly varying function compared to the convolution spectra of $J(\omega)$ which shift and broaden indefinitely toward lower energy as n increases (see Fig. 4.5). For this reason, the higher-order spectra no longer shift indefinitely towards lower energy. Instead, they converge to the spectral shape of the ordinary luminescence as shown schematically in Fig. 11.5(b). It should also be noted that the approximation by a differential operator cannot be applied to that part of the absorption spectrum which varies significantly while ω changes by the phonon energy ω_v . In a rigorous treatment of the linear electron-phonon interaction with a harmonic lattice, Kotani⁵² showed that in the mathematical limit of vanishing γ the zero-phonon line makes a predominant contribution to the Raman scattering of all orders. With a more elaborate argument, one finds that the above description which considers only the smooth part of the absorption spectra holds under the condition.

$$(\gamma/\omega_v)^{1/2} > S_T \exp(-S_T), \quad (11.4.38)$$

where ω_v represents the typical phonon energy. Putting realistic values: $\gamma \sim 10^8 \text{ s}^{-1}$ and $\omega_v \sim 10^{13} \text{ s}^{-1}$, one finds the condition $S_T > 7$. It is valid for most of the color centers with well-localized electrons (see also Fig. 10.5), but not for shallow centers in semiconductors (most of which, however, are to be classified among the weak coupling cases).

That the higher-order Raman processes make contributions comparable to those from lower-order ones irrespective of the coupling strength is an important feature of the resonance Raman scattering. One can see from Fig. 11.4(b) and Fig. 11.5(b) that increasingly higher order terms, with less memory of the incident energy Ω_1 , make predominant contributions as γ tends to 0 (note the limitation (11.4.38) in the strong coupling case). However, the way the Raman spectra of successive orders retain the memory and how they approach the ordinary luminescence depends very much upon the coupling strength or the modulation rapidity, as shown in Fig. 11.4(a) and Fig. 11.5(a) in the two contrasting cases.

11.5 Correlated absorption and emission of a localized electron

Let us now study the condition under which the resonance Raman scattering can be considered as luminescence in the conventional sense of the concept. Transforming the integration variables (τ, τ', σ) of (11.4.2) into $(\tau_0 \equiv (\tau + \tau')/2, \mu, \sigma)$ with the use of (11.3.5), we can write the scattering rate as

$$\begin{aligned} S(\Omega_1, \Omega_2) &\sim |(P_2)_{\text{eg}}|^2 \int_0^\infty d\tau_0 \int_{-\infty}^\infty d\mu \\ &\times \text{Tr}\{\exp[i(H_L + \Omega_2)\mu/2] \exp[-i(H_L + E_a + \delta V)\mu/2] A(\Omega_1, \tau_0) \\ &\times \exp[-i(H_L + E_a + \delta V)\mu/2] \exp[i(H_L + \Omega_2)\mu/2]\}, \end{aligned} \quad (11.5.1)$$

$$\begin{aligned} A(\Omega_1 \tau_0) &\equiv \exp[-i(H_L + E_a + \delta V - i\gamma)\tau_0] A(\Omega_1) \\ &\times \exp[i(H_L + E_a + \delta V + i\gamma)\tau_0], \end{aligned} \quad (11.5.2)$$

$$\begin{aligned} A(\Omega_1) &\equiv (P_1)_{\text{eg}}|^2 \int_{-\infty}^\infty d\sigma \exp[-i(H_L + E_a + \delta V)\sigma/2] \exp[-i(H_L + \Omega_1)\sigma/2] \\ &\times \rho_g \exp[-i(H_L + \Omega_1)\sigma/2] \exp[i(H_L + E_a + \delta V)\sigma/2]. \end{aligned} \quad (11.5.3)$$

According to (11.4.6), the trace of $A(\Omega_1)$ gives the absorption spectra:

$$\text{Tr}[A(\Omega_1)] = |(P_1)_{\text{eg}}|^2 F_a(\Omega_1) > 0. \quad (11.5.4)$$

The μ -integral in (11.5.1) has a structure similar to the σ -integral in (11.5.3) except for the interchange $e \rightleftharpoons g$ and $+\Omega_2 \rightleftharpoons -\Omega_1$, if $A(\Omega_1, \tau_0)$ is interpreted as a density matrix within the subspace e . In fact, $A(\Omega, 0) = A(\Omega)$ is such, with the positive normalization constant as given by (11.5.4). Then one can interpret $A(\Omega_1)$ as the density matrix immediately after the excitation with a photon Ω_1 , which evolves according to (11.5.2) as time elapses. Hence, the τ_0 -integrand in (11.5.1) can be interpreted as the emission spectra at time τ_0 after the optical excitation (time-resolved spectra). The absorption and emission are correlated through the propagator: $\exp[-i(H_L + E_a + \delta V - i\gamma)\tau_0]$.

The above interpretation holds only when τ_0 is much larger than the convergence range of σ - and μ -integrations in (11.5.3, 11.5.1). Since the latter is of the order of the reciprocal of the half-width h of the absorption and emission bands, we have to assume

$$\gamma \ll h \quad (11.5.5)$$

in order that the above interpretation holds for the dominant range of τ_0 .

Let us consider the local excitation represented by the density matrix ρ' . The operator $A(\Omega_1)/\text{Tr}[A(\Omega_1)]$ (see (11.5.3)) is an example of such a ρ' . We now

assume the ergodicity:

$$\begin{aligned}\rho(\tau) &\equiv \tau^{-1} \int_0^\tau d\tau' \exp[-i(H_L + E_a + \delta V)\tau'] \rho' \exp[i(H_L + E_a + \delta V)\tau'] \\ &\rightarrow \rho_e \quad (\tau \gg \tau_c \sim \omega_v^{-1}).\end{aligned}\quad (11.5.6)$$

Namely, starting from any locally excited state ρ' , the system approaches, with the relaxation time τ_c , always to the same thermal equilibrium with density matrix given by

$$\rho_e = \exp[-\beta(H_L + E_a + \delta V)] / \text{Tr}\{\exp[-\beta(H_L + E_a + \delta V)]\} \quad (11.5.7)$$

(compare with (11.4.3)). We expect this behavior simply because the system of lattice vibrations is very large (playing the role of heat bath in thermodynamic terminology), being insensitive to the way that the local excitation was introduced. Assuming $\gamma \ll \tau_c^{-1}$, the τ_0 -integral of $A(\Omega_1, \tau_0)$ in (11.5.1) can be carried out, with the use of eq. (11.5.6), the partial integration:

$$\begin{aligned}&\int_0^\infty d\tau \exp[-i(H_L + E_a + \delta V)\tau] \rho' \exp[i(H_L + E_a + \delta V)\tau] \exp(-2\gamma\tau) \\ &= \int_0^\infty d\tau \cdot d[\tau\rho(\tau)]/d\tau \cdot \exp(-2\gamma\tau) \sim \rho(\infty)/2\gamma,\end{aligned}$$

and eqs. (11.5.1) to (11.5.4). One finally obtains

$$S(\Omega_1, \Omega_2) \sim (2\gamma)^{-1} |(P_1)_{\text{eg}}|^2 F_a(\Omega_1) |(P_2)_{\text{ge}}|^2 F_e(\Omega_2) (\gamma \ll \tau_c^{-1}, h) \quad (11.5.8)$$

where $F_e(\Omega_2)$ is the normalized spectrum of emission from the relaxed excited state ρ_e :

$$F_e(\Omega_2) = \int_{-\infty}^{\infty} d\mu \exp[i\Omega_2\mu] \text{Tr}\{\exp[iH_L\mu] \exp[-i(H_L + E_a + \delta V)\mu] \rho_e\}. \quad (11.5.9)$$

Equation (11.5.8) represents the ideal situation where the resonance light scattering is factorized into two independent real processes, absorption and emission. The condition for this is shown in the parentheses. This factorization has been proved for two special cases: excitation of a localized electron with strong but almost linear interaction with the harmonic lattice of an insulator,³⁴ and core electron excitation interacting with conduction electrons in a metal.⁵³ We have shown that any ergodic system behaves in the same way.

There are many realistic systems, however, for which the ergodicity in its simplest form (11.5.6) is not valid within the radiative lifetime γ^{-1} . A typical situation may be that some modes of the system relax rapidly ($\tau_c \ll \gamma^{-1}$) but other modes rather slowly ($\tau' > \gamma^{-1}$). For example, in the case of a localized electron with a number of excited states designated by $\alpha = 1, 2, \dots$, the lattice will relax rapidly within

each adiabatic potential α , but the thermal equilibrium among different α s may be attained much more slowly through multiphonon processes. The parameter α may also represent branches of Jahn-Teller-split adiabatic potentials of degenerate electronic states, spin directions in a weakly spin-orbit coupled system, different sites of excitation in the problem of energy transfer, or the lattice sites of a self-trapped exciton in the problem of host excitation. The generalization of (11.5.6), with damping factor included, can then be written as

$$\begin{aligned} & \exp[-i(H_L + E_a + \delta V - i\gamma)t]\rho' \exp[i(H_L + E_a + \delta V + i\gamma)t] \\ & \rightarrow \sum_{\alpha} a_{\alpha}(t)\rho_{\alpha} \quad (\tau_c \ll t \sim O(\tau')), \end{aligned} \quad (11.5.10)$$

where ρ_{α} represents local equilibrium. To be precise, the l.h.s. should be understood as the time average over a time interval Δt ($\tau_c \ll \Delta t \ll \tau'$) around t . The time-dependent coefficients $a_{\alpha}(t)$ in (11.5.10) would satisfy the rate equations (with radiative damping rate γ_{α} and transition rate $p_{\alpha\neq\alpha'} \sim O(\tau'^{-1})$):

$$da_{\alpha}/dt = \sum_{\alpha'} p_{\alpha\alpha'} a_{\alpha'} - \gamma_{\alpha} a_{\alpha} \quad \left(\sum_{\alpha'} p_{\alpha\alpha'} = 0 \right), \quad (11.5.11)$$

independently from ρ' , but their initial values $a_{\alpha}(0)$ (at t such that $\tau_c < t \ll \tau'$) depend of course on the starting ρ' . For normalized ρ_{α} s, we have

$$\sum_{\alpha} a_{\alpha}(0) = 1 \quad (11.5.12)$$

since $\gamma\tau_c \ll 1$.

Let us now apply the generalized ergodicity Ansatz (11.5.10) to the time-resolved emission spectra after pulse excitation. For this purpose, let us first calculate the second-order response (11.3.3) to the step form excitation, at finite but large t ($\ll h^{-1}$) by changing the time variables into σ, μ , $\tau_0 = (\tau + \tau')/2$ and t_2 (see (11.3.5) or Fig. 11.2). Since the important range of σ and μ are $\sim O(h^{-1})$ and $\ll \gamma^{-1}$ according to the assumption (11.5.5), we can translate the graph invariantly by keeping the relative times σ, μ and τ_0 fixed but sliding t_2 between $\tau_0 + O(h^{-1})$ and $t + O(h^{-1})$. In this way, we obtain the response $\int_0^t d\tau_0(t - \tau_0) \cdots$ where the omitted part of the integrand is the same as the τ_0 -integrand of (11.5.1). In order to obtain the time-resolved emission spectrum L for unit pulse excitation, we have only to take the second derivative of the above expression; namely,

$$\begin{aligned} L(\Omega_1, \Omega_2, t) &= |(P_2)_{\text{ge}}|^2 \int_{-\infty}^{\infty} d\mu \text{Tr} \{ \exp[i(H_L + \Omega_2)\mu/2] \\ &\quad \times \exp[-i(H_L + E_a + \delta V)\mu/2] A(\Omega_1, t) \\ &\quad \times \exp[-i(H_L + E_a + \delta V)\mu/2] \\ &\quad \times \exp[i(H_L + \Omega_2)\mu/2] \}. \end{aligned} \quad (11.5.13)$$

This is in accordance with what we mentioned below eq. (11.5.4), as it should be. Inserting (11.5.10) into (11.5.13), and making use of the normalization (11.5.4), we obtain

$$L(\Omega_1, \Omega_2, t) = F_a(\Omega_1) \sum_{\alpha} a_{\alpha}(\Omega_1, t) F_e^{(\alpha)}(\Omega_2), \quad (11.5.14)$$

where the component emission $F_e^{(\alpha)}(\Omega_2)$ is given by (11.5.9) with ρ_e replaced by ρ_{α} . The coefficients $a_{\alpha}(t)$ depend on Ω_1 through the initial density matrix $A(\Omega_1)/\text{Tr}[A(\Omega_1)]$ and, hence, correlate the emission with the absorption spectra.

In the case of absorption spectra consisting of several component bands α which are separable in energy or polarization and of which each is adiabatically contingent to the relaxed excited state ρ_{α} , it is not unusual that the initial value $a_{\alpha}(\Omega_1, 0) = 1$ or 0 according as Ω_1 (with its polarization) belongs to the α -component band, $F_a^{(\alpha)}(\Omega_1)$, or not. Then we can write (11.5.14) as

$$L(\Omega_1, \Omega_2, t) = \sum_{\alpha\alpha'} F_a^{(\alpha')}(\Omega_1) a_{\alpha}^{(\alpha')}(t) F_e^{(\alpha)}(\Omega_2), \quad (11.5.15)$$

where $a_{\alpha}^{(\alpha')}(t)$ represents the solution of (11.5.11) with initial condition $a_{\alpha}^{(\alpha')}(0) = \delta_{\alpha\alpha'}$. In analyzing the so-called excitation spectra, the correlator $a_{\alpha}^{(\alpha')}(t)$ plays an important role.

The inequality $t \gg \tau_c$ characterizing the kinetic stage (where the kinetic equations (11.5.11) hold), on the one hand, and the inequality $t \gg h^{-1}$ assumed in the derivation of (11.5.13), on the other hand, are independent conditions. If $\tau_c \gg h^{-1}$ as is the case for a localized electron interacting strongly with the lattice (where τ_c^{-1} is of the order of the typical phonon energy ω_v), one can in principle observe the time-resolved emission spectra not only in the kinetic stage but even in the dynamical stage, by choosing the pulse excitation with duration time $\Delta t \ll \tau_c$ (so that one can follow the response for the interval $0 < t < \tau_c$) and the spectral width $\Delta\Omega_1 \ll h$ (so that one can study the Ω_1 dependence of the response). This problem will be treated more elaborately below. It should be noted, however, that the resolving power in time-resolved spectroscopy is limited by $\Delta\Omega_1 \sim \Delta\Omega_2 \sim \Delta t^{-1} \gg \tau_c^{-1}$ where Δt is the pulse duration or the observation interval.

As mentioned above, the correlation between absorption and emission is governed by the relaxation and thermal equilibrium state of the system within the excited-state subspace $|e\rangle$. To distinguish them from those in the ground-state subspace $|g\rangle$ so far described, we introduce the following notations:

$$\delta V^{(e)}(t) \equiv \exp[i(H_L + \delta V)t] \delta V \exp[-i(H_L + \delta V)t], \quad \langle\langle \delta V \rangle\rangle_e \equiv \text{Tr}[\rho_e \delta V],$$

$$\delta V(t) \equiv \exp[iH_L t] \delta V \exp[-iH_L t], \quad \langle\langle \delta V \rangle\rangle \equiv \text{Tr}[\rho_g \delta V]. \quad (11.5.16)$$

For instance, $\langle\langle\delta V^{(e)}(t)\rangle\rangle$ starts from $\langle\langle\delta V\rangle\rangle = 0$ (see (11.4.10)) at $t = 0$, and tends to $\langle\langle\delta V\rangle\rangle_e$ at $t = \infty$ under the ergodicity assumption. In fact, in the expression

$$\begin{aligned}\langle\langle\delta V^{(e)}(\tau')\rangle\rangle &\equiv \text{Tr}\{\rho_g \exp[iH_L + \delta V]\tau'\} \delta V \exp[-i(H_L + \delta V)\tau']\} \\ &= \text{Tr}\{\exp[-i(H_L + \delta V)\tau']\rho_g \exp[i(H_L + \delta V)\tau']\delta V\}\end{aligned}$$

obtained by cyclic permutation of operators allowed within $\text{Tr}\{\cdots\}$, we take the long-time average, $\tau^{-1} \int_0^\tau d\tau'$, of both sides and note that the operator in the bracket $\{\cdots\}$ on the r.h.s. can be replaced by ρ_e for arbitrary ρ' (which in the present case is ρ_g). Then we obtain

$$\text{Lim}_{\tau \rightarrow \infty} \left[\tau^{-1} \int_0^\tau d\tau' \langle\langle\delta V^{(e)}(\tau')\rangle\rangle \right] = \text{Lim}_{\tau \rightarrow \infty} \langle\langle\delta V^{(e)}(\tau)\rangle\rangle = \text{Tr}[\rho_e \delta V] \quad (11.5.17)$$

as mentioned above.

With this in mind, we define the instantaneous fluctuation by

$$\delta V_e(t) \equiv \delta V^{(e)}(t) - \langle\langle\delta V^{(e)}(t)\rangle\rangle, \quad (11.5.18)$$

and rewrite (11.4.15) as

$$\begin{aligned}&\langle\langle \exp[i(H_L + \delta V)\tau'] \exp[iH_L\mu] \exp[-i(H_L + \delta V)_\tau] \exp[-iH_L\sigma] \rangle\rangle \\ &= \langle\langle S_\mu^{(e)}(\tau') S_\sigma^\dagger \rangle\rangle.\end{aligned} \quad (11.5.19)$$

In expanding S_σ defined by (11.4.5), we follow the evolution within the excited-state subspace in conformity with the above; namely, we start from the differential equation

$$(d/d\sigma)S_\sigma = -i \exp[iH_L\sigma] \delta V \exp[-i(H_L + \delta V)\sigma] = -i S_\sigma \delta V^{(e)}(\sigma) \quad (11.5.20)$$

(note the difference of the r.h.s from that of (11.4.7)), and repeatedly integrate it to have

$$\begin{aligned}S_\sigma &= 1 - i \int_0^\sigma d\sigma_1 S_{\sigma_1} \delta V^{(e)}(\sigma_1) \\ &= 1 - i \int_0^\sigma d\sigma_1 \delta V^{(e)}(\sigma_1) - \int_0^\sigma d\sigma_1 \int_0^{\sigma_1} d\sigma_2 \delta V^{(e)}(\sigma_2) \delta V^{(e)}(\sigma_1) + \cdots.\end{aligned} \quad (11.5.21)$$

One can thus expand (11.5.19) as

$$\begin{aligned}\langle\langle S_\mu^{(e)}(\tau') S_\sigma^\dagger \rangle\rangle &= 1 \\ &\quad - i \int_0^\mu d\mu_1 \langle\langle \delta V^{(e)}(\mu_1 + \tau') \rangle\rangle + i \int_0^\sigma d\sigma_1 \langle\langle \delta V^{(e)}(\sigma_1) \rangle\rangle \\ &\quad + \int_0^\mu d\mu_1 \int_0^\sigma d\sigma_1 \langle\langle \delta V^{(e)}(\mu_1 + \tau') \delta V^{(e)}(\sigma_1) \rangle\rangle\end{aligned}$$

$$\begin{aligned}
& - \int_0^\mu d\mu_1 \int_0^{\mu_1} d\mu_2 \langle\langle \delta V^{(e)}(\mu_2 + \tau') \delta V^{(e)}(\mu_1 + \tau') \rangle\rangle \\
& - \int_0^\sigma d\sigma_1 \int_0^{\sigma_1} d\sigma_2 \langle\langle \delta V^{(e)}(\sigma_1) \delta V^{(e)}(\sigma_2) \rangle\rangle + (\text{higher-order terms}).
\end{aligned} \tag{11.5.22}$$

In calculating the logarithm of (11.5.22), one finds that the second-order term of the second line of the r.h.s. can be added to the first-order term of the third to fifth lines resulting in the change of $\delta V^{(e)}(\sigma_1)$ into $\delta V_e(\sigma_1)$, etc. in accordance with (11.5.18). Thus we have

$$\begin{aligned}
\ln \langle\langle S_\mu^{(e)}(\tau') S_\sigma^\dagger \rangle\rangle = & \left[-i \int_0^\mu d\mu_1 \langle\langle \delta V^{(e)}(\tau' + \mu_1) \rangle\rangle \right. \\
& - \int_0^\mu d\mu_2 \int_0^{\mu_2} d\mu_1 \langle\langle \delta V^{(e)}(\tau' + \mu_1) \delta V^{(e)}(\tau' + \mu_2) \rangle\rangle + \cdots \left. \right] \\
& + \left[i \int_0^\sigma d\sigma_1 \langle\langle \delta V^{(e)}(\sigma_1) \rangle\rangle \right. \\
& - \int_0^\sigma d\sigma_1 \int_0^{\sigma_1} d\sigma_2 \langle\langle \delta V_e(\sigma_1) \delta V_e(\sigma_2) \rangle\rangle + \cdots \left. \right] \\
& + \left[\int_0^\mu d\mu_1 \int_0^\sigma d\sigma_1 \langle\langle \delta V_e(\tau' + \mu_1) \delta V_e(\sigma_1) \rangle\rangle + \cdots \right] \tag{11.5.23}
\end{aligned}$$

where we have rearranged the terms into μ -integrals, σ -integrals and mixed integrals. Since the first and second brackets represent the cumulant expansion of the logarithms of $\langle\langle S_\mu^{(e)}(\tau') \rangle\rangle$ and $\langle\langle S_\sigma^\dagger \rangle\rangle$, respectively, one can rewrite (11.5.23), with the use of K_e defined by (11.4.17) with $\delta V(t)$ replaced by $\delta V_e(t)$, as

$$\begin{aligned}
\langle\langle S_\mu^{(e)}(\tau') S_\sigma^\dagger \rangle\rangle = & \langle\langle S_\mu^{(e)}(\tau') \rangle\rangle \langle\langle S_\sigma^\dagger \rangle\rangle \\
& \times \exp[K_e(-\sigma, -\mu, -\tau') + \cdots]. \tag{11.5.24}
\end{aligned}$$

The higher-order cumulants shown by \cdots in eqs. (11.5.23, 11.5.24) vanish for the linear interaction with a harmonic lattice as was mentioned in the preceding section.

Equations (11.4.16) and (11.5.24) represent the dual nature of the resonant secondary radiation. In (11.4.16), the first two factors represent the polarizability (when Fourier-transformed over the half-space $\tau > 0$) while the third factor represents how the polarizability is correlated with itself at different time μ . In (11.5.24), the second factor represents the absorption (when Fourier-transformed over $-\infty < \sigma < +\infty$) as is seen from (11.4.6), the first factor the emission at time τ' after the absorption, while the third factor describes how the preceding two are correlated. It should be noted, however, that in spite of extracting this correlation factor, the first factor keeps its memory of the initial state though the averaging $\langle\langle \cdots \rangle\rangle$ on ρ_g , and depends on τ' (note that $[H_L, \delta V] \neq 0$), in contrast to (11.4.16) where the second factor on

the r.h.s. no longer depends on μ . Moreover, when we take σ , μ and τ' as independent variables, the ranges of integration for $S(\Omega_1, \Omega_2)$ are no longer mutually independent because of the restriction that $\tau = \tau' - \sigma + \mu \geq 0$. The duality is not symmetric.

It is instructive to study this second picture for the resonant secondary radiation in the two limits of weak and strong coupling. In the weak coupling case, the difference in propagation through $H_L + \delta V$ and H_L is not important. Replacing $\delta V^{(e)}(t)$ and $\delta V_e(t)$ by $\delta V(t)$ in (11.5.23), and making use of $\langle \delta V(t) \delta V(t') \rangle = 2\Gamma \delta(t - t')$, one can write

$$\ln \langle S_\mu(\tau') S_\sigma^\dagger \rangle = -\Gamma(|\mu| + |\sigma|) + 2\Gamma \int_0^\mu d\mu_1 \int_0^\sigma d\sigma_1 \delta(\mu_1 + \tau' - \sigma_1). \quad (11.5.25)$$

Inserting this in (11.4.2), we obtain the well-known result^{32,35}

$$S(\Omega_1, \Omega_2) = 2\pi |(P_1)_{\text{eg}}|^2 |(P_2)_{\text{ge}}|^2 [(\Omega_1 - E_a)^2 + (\Gamma + \gamma)^2]^{-1} \\ \times \{ \delta(\Omega_2 - \Omega_1) + (\Gamma/\gamma)(\Gamma + \gamma)/\pi [(\Omega_2 - E_a)^2 + (\Gamma + \gamma)^2]^{-1} \}, \quad (11.5.26)$$

which is also in agreement with (11.4.34). It should be noted here that the last term in (11.5.25), originating from the correlation term K_e in (11.5.24), is non-vanishing only for those diagrams such that the intervals (t_1, t'_1) and (t_2, t'_2) are overlapping, namely such that absorption and emission are not separable in time ordering. Only these diagrams (see Figs. 11.2 and 11.3) contribute to the first term in $\{\cdot \cdot \cdot\}$ of (11.5.26).

In the strong coupling case, there is a significant difference in propagation between H_L and $H_L + \delta V$ which causes the time-dependent emission as mentioned before. The expression (11.5.13) for the time-dependent emission spectra for a pulse excitation can be rewritten, with the use of S_σ defined by (11.4.5), as

$$L(\Omega_1, \Omega_2, t) = |(P_2)_{\text{ge}}|^2 \int_{-\infty}^{\infty} d\mu \int_{-\infty}^{\infty} d\sigma \exp[-i(\Omega_1 - E_a)\sigma \\ + i(\Omega_2 - E_a)\mu - 2\gamma t] \langle S_\mu^{(e)}(t + \{\sigma - \mu\}/2) S_\sigma^\dagger \rangle, \quad (11.5.27)$$

where we have made use of the symmetrized time $t \equiv (\tau + \tau')/2$ for the elapsed time between absorption and emission. The correlation function in (11.5.27) is explicitly given by (11.5.23) where $\tau' = t + (\sigma - \mu)/2$ is to be inserted. It would be enough to expand it in powers of σ and μ up to the quadratic terms considering the strong coupling condition. Making use of the relation such as

$$-i \langle \langle d/dt \delta V^{(e)}(t) \rangle \rangle = \langle \langle [H_L + \delta V, \delta V^{(e)}(t)] \rangle \rangle = \langle \langle \delta V, \delta V_e(t) \rangle \rangle,$$

one obtains

$$\begin{aligned} \ln \langle \langle S_\mu^{(e)}(t + \{\sigma - \mu\}/2) S_\sigma^\dagger \rangle \rangle &= [-\langle \langle (\delta V)^2 \rangle \rangle \sigma^2/2] + [-i \langle \langle \delta V^{(e)}(t) \rangle \rangle \mu \\ &\quad - \langle \langle (\delta V_e(t))^2 \rangle \rangle \mu^2/2] + [\langle \langle \{\delta V, \delta V_e(t)\} \rangle \rangle \sigma \mu] \end{aligned} \quad (11.5.28)$$

where $\{A, B\} \equiv (AB + BA)/2$ is the symmetrized product.

But for the last term on the r.h.s of (11.5.28), the Fourier transform (11.5.27) would give the product of the absorption spectrum with its peak at E_a and dispersion $\langle \langle (\delta V)^2 \rangle \rangle = D^2$ and the time-dependent emission spectrum with its peak at $E_a + \langle \langle \delta V^{(e)}(t) \rangle \rangle$ and dispersion $\langle \langle (\delta V_e(t))^2 \rangle \rangle = \langle \langle [\delta V^{(e)}(t) - \langle \langle \delta V^{(e)}(t) \rangle \rangle]^2 \rangle \rangle$ which reduces to the former at $t = 0$. In fact, however, the emission is correlated with the absorption through this last term which is the $\sigma - \mu$ cross term. To see this, let us consider the incident photon flux of spectral width $\delta : I(\Omega_1) = I_0(2\pi\delta^2)^{-1/2} \times \exp[-(\Omega_1 - \Omega_0)^2/2\delta^2]$. In order that time-resolved emission spectroscopy be useful for the study of a relaxation process which takes place within time $\sim \tau_c$, the duration time Δt of the pulse excitation as well as that of the observation should be $\ll \tau_c$. The spectral widths δ for excitation and $\Delta\Omega_2$ (resolving power) for observation should be much larger than τ_c^{-1} (but can still be made smaller than D). The time-resolved emission spectrum for this pulse excitation is given by $\int d\Omega_1 L(\Omega_1, \Omega_2, t) \times I(\Omega_1)$. Making use of (11.5.27, 11.5.28), one finds that the peak and dispersion of the emission spectrum are given as follows:

$$\overline{\Omega_2(t)} = \langle \langle \delta V^{(e)}(t) \rangle \rangle + \langle \langle \{\delta V, \delta V_e(t)\} \rangle \rangle [\langle \langle (\delta V)^2 \rangle \rangle + \delta^2]^{-1} (\Omega_0 - E_a), \quad (11.5.29)$$

$$\overline{\{\Omega_2(t) - \overline{\Omega_2(t)}\}^2} = \langle \langle (\delta V_e(t))^2 \rangle \rangle - \langle \langle \{\delta V, \delta V_e(t)\} \rangle \rangle^2 [\langle \langle (\delta V)^2 \rangle \rangle + \delta^2]^{-1}. \quad (11.5.30)$$

One can see how the correlation effect mentioned above disappears for white incidence ($\delta \gg D$). For the general case, and at $t = 0$, the peak of the emission band is at the weighted mean $(D^2\Omega_0 + \delta^2 E_a)/(D^2 + \delta^2)$ of Ω_0 and E_a while the width is given by $D\delta/(D^2 + \delta^2)^{1/2}$ which is closer to the smaller of D and δ when they are much different. For $t \gg \tau_c$, the second terms of (11.5.29, 11.5.30) disappear (corresponding to the loss of memory of the excitation conditions, Ω_0 and δ) while the first term approaches the values $E_a + \langle \langle \delta V \rangle \rangle_e \equiv E_e$ and $\langle \langle (\delta V - \langle \langle \delta V \rangle \rangle_e)^2 \rangle \rangle_e$. The average $-\langle \langle \delta V \rangle \rangle_e (>0)$ is the so-called Stokes shift.

Calculated time-resolved emission spectra for the nearly monochromatic and nearly white excitations are shown schematically in Fig. 11.6(b) by solid and thick dashed lines, respectively, for a localized electron with linear interaction with the harmonic lattice, corresponding to the relaxation within the excited state as depicted

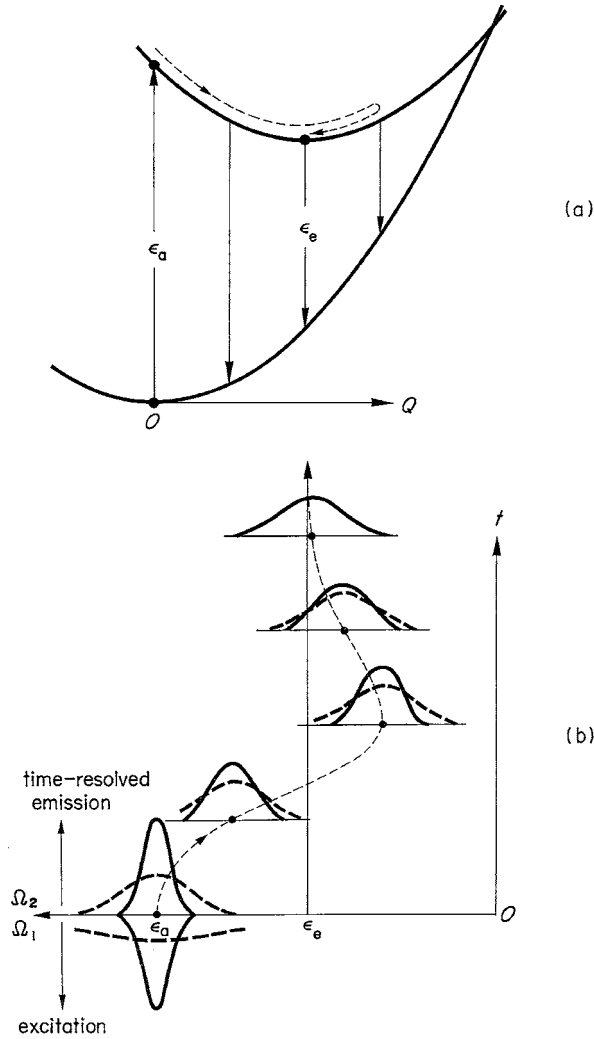


Fig. 11.6 (a) Configuration coordinate model, and (b) time-resolved emission spectra, of a strong coupled localized electron-phonon system under photoexcitation.⁴⁵

by the configuration coordinate model in Fig. 11.6(a). It is left to the reader to derive these results from (11.5.29, 11.5.30), with the help of the description in Section 4.4. Note that Fig. 4.3(b) corresponds to the nearly white excitation (thick dashed lines) of Fig. 11.6(b).

In addition to the stationary and time-resolved spectra for the second order optical process described in the preceding and the present section, respectively, another type of spectrum described in Appendix 2 turns out to be useful for some purposes.

11.6 Resonant light scattering in an exciton-phonon system

Let us now consider the resonant excitation of host electrons instead of a localized electron, namely an exciton as the resonant intermediate state of light scattering. This is a typical case in which the radiative damping γ is an operator with singular dependence $N\delta_{\mathbf{K}\mathbf{0}}$ on the exciton wave vector \mathbf{K} where N is the number of unit cells in the macroscopic crystal. Radiative damping takes place only when the exciton with $\mathbf{K} = \mathbf{0}$ is created or annihilated through the transition dipole \mathbf{P}_{eg} which has singularity of the form $N^{1/2}\delta_{\mathbf{K}\mathbf{0}}$ as was mentioned in Section 8.4. Confining ourselves to the lowest exciton band and to a single mode of lattice vibrations representatively with their respective suffixes λ and s being omitted, we can write the excited state Hamiltonian in (11.4.1) as

$$\begin{aligned} H_e &= \sum_{\mathbf{K}} \sum_{\mathbf{K}'} |\mathbf{K} \rangle [(E(\mathbf{K}) + H_L)\delta_{\mathbf{K}\mathbf{K}'} + (H_I)_{\mathbf{K}\mathbf{K}'}] \langle \mathbf{K}' | \\ &\equiv H_{(e)} + H_L + H_I \equiv H_0 + H_I, \end{aligned} \quad (11.6.1)$$

with the use of linear interaction H_I of an exciton with a single mode of phonons (different modes will be considered later) given by

$$(H_I)_{\mathbf{K}\mathbf{K}'} = V(\mathbf{K} - \mathbf{K}')b_{\mathbf{K}-\mathbf{K}'} + V^*(-\mathbf{K} + \mathbf{K}')b_{-\mathbf{K}+\mathbf{K}'}^\dagger \quad (11.6.2)$$

(see (10.1.1)) and the cross-section for light scattering, (11.3.1), as

$$\begin{aligned} S(\Omega_1, \Omega_2) &= |(P_1)_{\text{eg}}|^2 |(P_2)_{\text{ge}}|^2 \\ &\times \sum_m w_m \sum_n |\langle \mathbf{0}n | [\Omega_1 + E_m - H_e]^{-1} | \mathbf{0}m \rangle|^2 \delta(\omega_{nm} - \Omega_1 + \Omega_2) \end{aligned} \quad (11.6.3)$$

where E_m and E_n are the initial and final vibrational states of H_L . Since the $|\ell\rangle$ s in (11.3.1) represent *all* the eigenstates of H_e , namely

$$H_e = \sum_{\ell} |\ell\rangle E_{\ell} \langle \ell|,$$

we have made use of the abbreviation for the ℓ -summation in (11.3.1) in operator form

$$\sum_{\ell} |\ell\rangle [\Omega_1 - \omega_{\ell m}]^{-1} \langle \ell| = [\Omega_1 + E_m - H_e]^{-1} \equiv [z - H_e]^{-1} \quad (11.6.4)$$

in deriving (11.6.3).

We can expand the resolvent (11.6.4) in a power series in H_I , as we did in Section 8.7, as

$$\begin{aligned} [z - H_e]^{-1} &= [z - H_0]^{-1} \\ &+ [z - H_0]^{-1} H_I [z - H_0]^{-1} \\ &+ [z - H_0]^{-1} H_I [z - H_0]^{-1} H_I [z - H_0]^{-1} + \dots, \end{aligned} \quad (11.6.5)$$

of which the n th order term inserted into (11.6.3) corresponds to the Raman process in which n phonons are emitted or absorbed. In the following, we will describe the cases of (i) a direct exciton (with its band bottom at $\mathbf{K} = \mathbf{0}$) and (ii) an indirect exciton (band bottom at $\mathbf{K} \neq \mathbf{0}$) separately since the situations are quite different.

11.6.1 Direct exciton

For simplicity, we confine ourselves to the absolute zero of temperature where only the emission of phonons takes place, and consider for the moment longitudinal optical phonons with dispersionless energy ω_ℓ . As for the excitonic part, the n th-order term gives a contribution

$$\begin{aligned} & \langle \mathbf{0} | [\Omega_1 - n\omega_\ell - E(\mathbf{0})]^{-1} \sum_{\mathbf{K}_1} \dots \sum_{\mathbf{K}_{n-1}} \mathbf{V}^*(-\mathbf{K}_{n-1}) \\ & \quad \times [\Omega_1 - (n-1)\omega_\ell - E(\mathbf{K}_{n-1})]^{-1} \mathbf{V}^* \dots [\Omega_1 - \omega_\ell - E(\mathbf{K}_1)]^{-1} \mathbf{V}^*(\mathbf{K}_1) \\ & \quad \times [\Omega_1 - E(\mathbf{0})]^{-1} |\mathbf{0}\rangle|^2 \delta(\Omega_2 + n\omega_\ell - \Omega_1) \end{aligned} \quad (11.6.6)$$

to the n th-order Raman Stokes line at $\Omega_2 = \Omega_1 - n\omega_\ell$ ($n \geq 2$). One can assume that the wave vector $\mathbf{K}_{n'}$ of the intermediate exciton state ($1 \leq n' \leq n-1$) with no radiative decay γ does not take the value $\mathbf{0}$ since such a case (with enormous γ -value ($\sim N$) as mentioned before) is already counted as the $n'(<n)$ th order term of (11.6.5).

As a function of Ω_1 , the cross-section $S(\Omega_1, \Omega_2)$ is enhanced at $\Omega_1 = E(\mathbf{0}) + n\omega_\ell$ ($n \geq 2$) due to the *outgoing* resonance (vanishing of the leftmost energy denominator $[\Omega_1 - n\omega_\ell - E(\mathbf{0})]$ in (11.6.6)), the scattered light being at $\Omega_2 = E(\mathbf{0})$. This remarkable feature was first found by Goto, Takahashi and Ueta⁵⁴ in their study, in CuBr, of the *excitation spectra* for free-exciton luminescence at $\Omega_2 = E(\mathbf{0})$ (rather than as the Raman spectra), with periodic structure $\Omega_1 = E(\mathbf{0}) + n\omega_\ell$ in the incident energy Ω_1 , as shown in Fig. 11. 7, and was explained in terms of multiple *real* scattering processes such that all the energy denominators in (11.6.6) vanish.

Of course, the n th-order term, (11.6.6), of $S(\Omega_1, \Omega_2)$ contains the factor $\delta(\Omega_2 + n\omega_\ell - \Omega_1)$ which indicates the existence of the Raman line at $\Omega_2 = \Omega_1 - n\omega_\ell$ with constant Stokes shift $n\omega_\ell$ from incident energy Ω_1 as a variable, and the entire features of coexistent Raman lines and luminescence band were extensively studied by Permogorov and others in a variety of ionic semiconductors.⁵⁵ Equation (11.6.6) indicates that as a function of Ω_1 the intensity of the n th-order Raman line is enhanced at $\Omega_1 = E(\mathbf{0})$ and $\Omega_1 = E(\mathbf{0}) + n\omega_\ell$ (namely at $\Omega_2 = E(\mathbf{0})$ where the Raman line is superposed on the broad luminescence band), the so-called *incoming* and *outgoing resonance* effects, respectively. Under the assumption of real transitions only (note that the exact formula (11.6.6) contains the virtual processes), the populations of

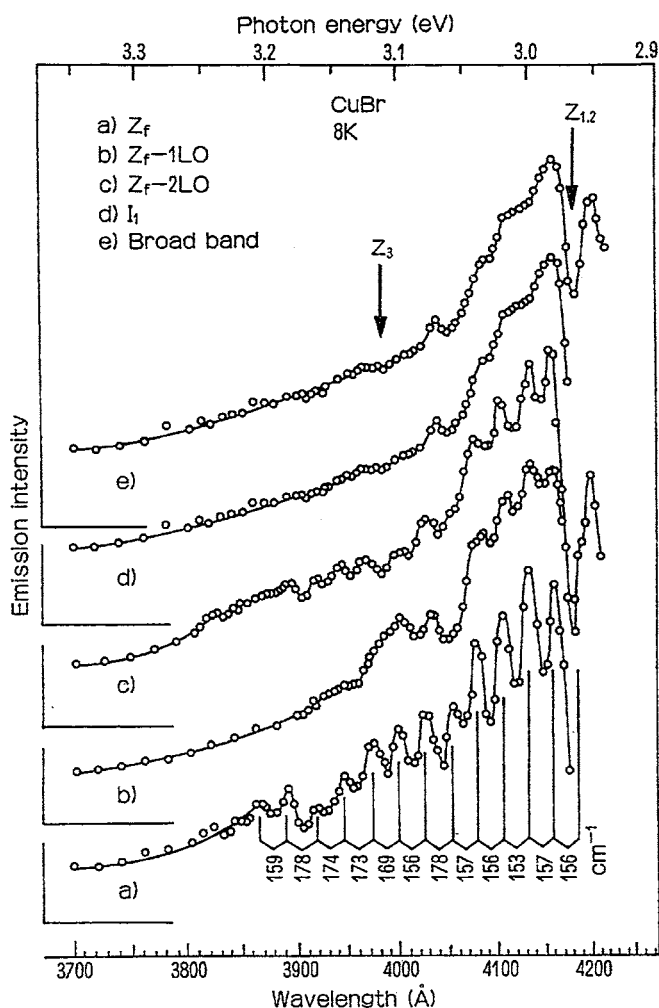


Fig. 11.7 Excitation spectra of free- and bound-exciton emissions and a broad band of pair recombination type in CuBr. Due to Goto *et al.*⁵⁴

the n 'th intermediate excitation states are approximately equal due to the balance of numbers of excitons cascading down into and out of the n 'th step, as described by Permogorov.

The entire behavior consisting of incoming, outgoing and *intermediate state* resonances was clearly observed in CdTe by A. Nakamura and Weisbuch.⁵⁶ Of particular interest is the fact that at resonances the sharp Raman lines become broader and coalesce with the emission band, instead of simply being superimposed upon the latter and keeping their own sharpness as had been found in other materials. They ascribed this to the long lifetime of excitons in their high-purity samples such that the excitons have enough time to be thermalized.

Kurita, Sakai and Kotani studied,⁵⁷ in their calculation of Raman spectra with acoustic-as well as optical-phonon scattering considered, how the virtual and real scattering processes compete and how multiple scattering by acoustic phonons makes a dominant contribution to the final stage of exciton thermalization towards the band bottom, resulting in the broad emission band.

In this connection, it is interesting to note that in the Raman process with one-acoustic- and one-optical-phonon emission in ZnTe, the former is emitted first in the incoming resonance and the latter first in the outgoing resonance because of the better *intermediate state resonance* in the respective cases, as noted by Oka and Cardona.⁵⁸

The Raman process with one-phonon emission can take place only when one takes into account the finite wave number of photons so far neglected. For a quantitative description of the resonance Raman scattering inclusive of higher-order-phonon processes, one has to resort to the polariton picture (see Section 8.6). According to this picture, the first-order Raman scattering *within matter* is nothing other than the (intra- or interbranch) scattering of a polariton by one phonon instead of being the second-order optical process. This improved description is essential for the resonant Brillouin scattering in which acoustic phonons are emitted or absorbed. As a result, more than a couple of Stokes and anti-Stokes Brillouin lines can appear, with the latter being more intense (contrary to the non-resonant Brillouin lines with ratio $(\bar{n} + 1) : \bar{n}$) and more shifted.⁵⁹ We refer to an extensive review article by Weisbuch and Ulbrich⁶⁰ for the polariton description of resonance light scattering via an exciton, where one can find also many useful tables of the physical parameters of exciton-polaritons in a number of ionic semiconductors.

Among distinguished works which appeared after this review article, we refer here to an extensive study of β -ZnP₂ mainly due to Arimoto, Sugisaki and K. Nakamura,^{61–65} but partly in collaboration with other groups. In spite of a complicated structure with eight molecules per unit cell and extremely anisotropic optical properties with singlet excitons for $\mathbf{E} \parallel c$ and triplet excitons for $\mathbf{E} \parallel b$ polarizations, both excitons consist of beautiful Bohr series (characteristic of spherical bands) up to 4th and 7th sharp lines, and exciton–phonon interactions are extremely weak except for one LO-phonon out of 72 modes. A variety of Raman lines with resonance enhancements interplaying with luminescence, and singlet–triplet conversion dynamics have been studied, revealing a detailed amount of what is taking place.

11.6.2 Indirect exciton

Resonance Raman spectroscopy has also proved to be a powerful tool for studying indirect excitons, for a reason somewhat different from the direct exciton case.

With the use of incident light resonant to the indirect exciton, one can pass through the radiatively singular $\mathbf{K} = \mathbf{0}$ exciton state merely as a non-resonant intermediate state and concentrate on the detailed study of intra- and intervalley scattering of indirect excitons, as shown by von der Osten and collaborators⁶⁶⁻⁷⁶ in their extensive studies on silver halides. We will describe here the main points of their results. Confining oneself to absolute zero of temperature, one has to replace, in the expression (11.6.3), the direct transition matrix element $(\mathbf{P}_1)_{\text{eg}}$ by the indirect one: $\sum_{\lambda'} (\Omega_1 - E_{\lambda'\mathbf{0}})^{-1} \mathbf{V}_{\lambda\lambda's}^*(-\mathbf{K})(\mathbf{P}_1)_{\lambda'g}$ and $(\mathbf{P}_2)_{\text{ge}}$ by $\sum_{\lambda'} (\mathbf{P}_2)_{g\lambda'} \mathbf{V}_{\lambda'\lambda s}(\mathbf{K}')(\Omega_2 - E_{\lambda'\mathbf{0}})^{-1}$ where $\sum_{\lambda'}$ is to be extended over internal states λ' of the direct exciton, and $(-\mathbf{K}, s)$ and (\mathbf{K}', s') are the wave vectors and the modes of the phonons emitted in the indirect excitation and de-excitation, respectively, of the electronic system. (This amounts to taking the third and lower lines in (11.6.5) and rewriting the leftmost and rightmost blocks of all terms with momentum-conserving phonons as shown above.) Confining oneself to the indirect excitons in the lowest internal state $\lambda (= 1s)$, one has to replace the ket vector $|\mathbf{0}\rangle$ by $|\lambda\mathbf{K}, 1_{-\mathbf{K},s}\rangle$ and bra vector $\langle \mathbf{0}|$ by $\langle \lambda\mathbf{K}', 1_{-\mathbf{K},s}, n|$ where n represents the state of phonons emitted while the indirect exciton is subject to intra- or intervalley scattering from \mathbf{K} to \mathbf{K}' . Putting $E_m = 0$ (no phonons in the initial state), omitting the index λ and denoting the energy of the intermediately emitted phonons by ω_n , one can write

$$\begin{aligned}
 S(\Omega_1, \Omega_2) = & \sum_{\lambda'} \sum_{\lambda''} |(\mathbf{P}_1)_{\lambda'g}|^2 |(\mathbf{P}_2)_{\lambda''e}|^2 (\Omega_1 - E_{\lambda'\mathbf{0}})^{-2} (\Omega_2 - E_{\lambda''\mathbf{0}})^{-2} \\
 & \times \sum_{s,K} \sum_{s',K'} |V_{\lambda\lambda's'}(\mathbf{K}') V_{\lambda\lambda's}^*(-\mathbf{K})|^2 \\
 & \times \sum_n |\langle \mathbf{K}', n | [\Omega_1 - \omega_{-\mathbf{K},s} - H_e]^{-1} | \mathbf{K} \rangle|^2 \\
 & \times \delta(\Omega_2 + \omega_{-\mathbf{K},s} + \omega_{\mathbf{K}',s'} + \omega_n - \Omega_1),
 \end{aligned} \tag{11.6.7}$$

where we have, consistently with the addition of the term $-\omega_{-\mathbf{K},s}$ in the energy denominator, omitted the phonon $(-\mathbf{K}, s)$ from H_e as well as from the bra vector $\langle |$ and the ket vector $| \rangle$ of the intermediate states, simply for the sake of avoiding unnecessarily complicated notation.

The resolvent in (11.6.7) can be expanded in the same way as in (11.6.5) where $z \equiv \Omega_1 - \omega_{-\mathbf{K},s}$ in the present case. Here, however, we take the *renormalized* expansion in which the unperturbed Hamiltonian H_0 is supplemented with the self-energy $\Sigma = \Delta + i\Gamma$ (diagonalized together with H_0) as we did in Section 10.2:

$$\begin{aligned}
 [z - H_e]^{-1} = & [z - H_0 - \Sigma]^{-1} \\
 & + [z - H_0 - \Sigma]^{-1} H_1 [z - H_0 - \Sigma]^{-1} \\
 & + [z - H_0 - \Sigma]^{-1} H_1 [z - H_0 - \Sigma]^{-1} H_1 [z - H_0 - \Sigma]^{-1} + \dots
 \end{aligned} \tag{11.6.8}$$

In the expansion (11.6.8) which holds exactly,⁷⁷ one should restrict the intermediate states in the product of each term so as to be different from each other and from the initial and final states, because those terms which have returned to the same intermediate states as before are factorized and already included within the self-energy Σ in each factor of each term in (11.6.8).

With the incident energy Ω_1 which is nearly resonant to the indirect absorption edge, the produced indirect exciton has \mathbf{K} not far from one of the equivalent band bottoms \mathbf{K}_α (the bottoms are at four equivalent L-points on the Brillouin zone edge in the (111)-directions in the case of silver halides), and its kinetic energy ε can be approximated by a quadratic function of momentum measured from the bottom, $\mathbf{k}_\alpha \equiv \mathbf{K} - \mathbf{K}_\alpha$,

$$\varepsilon \equiv E(\mathbf{K}) - E_b = \mathbf{k}_\alpha (1/2 \overline{\overline{M}}_\alpha) \mathbf{k}_\alpha \quad (11.6.9)$$

where $\overline{\overline{M}}_\alpha$ is the effective mass tensor of the i th valley and E_b is the common bottom energy. The \mathbf{K} -summation in (11.6.7) would diverge at

$$E(\mathbf{K}) = \Omega_1 - \omega_{-\mathbf{K},s} \quad (11.6.10)$$

but for the imaginary part $\Gamma_{\text{tot}}(\mathbf{K}) \equiv \Gamma_t(\varepsilon)$ of the self-energy which represents the scattering rate of the \mathbf{K} -exciton with kinetic energy ε due to all energetically possible processes of emitting phonons which in fact arises in the second and subsequent lines in (11.6.8) as will be described below. Under the assumption of small $\Gamma_t(\varepsilon)$, the \mathbf{K} -summation gives⁶⁶

$$\sum_{\mathbf{K}} \rightarrow \int_0 \varepsilon^{1/2} [(E_b + \varepsilon + \omega_{-\mathbf{K},s} - \Omega_1)^2 + \{\Gamma_t(\varepsilon)\}^2]^{-1} d\varepsilon \propto \varepsilon'^{1/2} / \Gamma_t(\varepsilon') \quad (11.6.11)$$

where \mathbf{K} in $\omega_{-\mathbf{K},s}$ is replaced by \mathbf{K}_α and then $\omega_{-\mathbf{K},s}$ by a constant ω_s independent of the equivalent α , and $\varepsilon' = \Omega_1 - E_b - \omega_s$ where (\cdots) in the denominator of the integrand vanishes.

With small kinetic energy ε , the only possible process contributing to $\Gamma_t(\varepsilon)$ is intravalley scattering with emission of a longitudinal acoustic (LA) phonon. Approximating the effective mass tensor $\overline{\overline{M}}_\alpha$ by isotropic mass M^* , and noting that the matrix element for exciton-acoustic-phonon scattering is proportional to deformation potential $\Xi_{\text{ex}} = \Xi_e - \Xi_h$ times the square root of the phonon wave vector q (see Sections 9.4 and 10.1), one finds that

$$\Gamma_{\text{ac}}(\varepsilon) \propto \Xi_{\text{ex}}^2 \sum_q q \delta[\{\mathbf{k}^2 - (\mathbf{k} - \mathbf{q})^2\} / 2M^* - c_\ell |q|] \quad (11.6.12)$$

where c_ℓ is the velocity of the longitudinal sound wave. It turns out to be given by

$$\Gamma_{\text{ac}}(\varepsilon) \propto \Xi_{\text{ex}}^2 \varepsilon \quad (11.6.13)$$

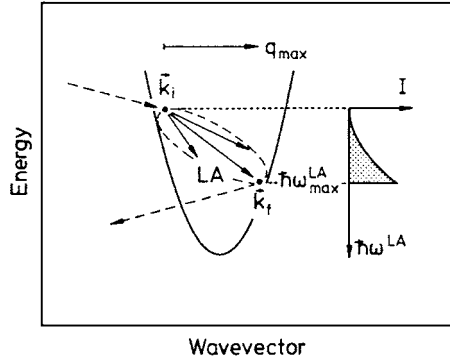


Fig. 11.8 Schematic diagram of exciton dispersion and intravalley scattering by long-wavelength LA phonons. The dashed arrows correspond to TO(L) phonons. The final states \mathbf{k}_f are along the dashed conical section. Also shown is the lineshape of the scattering spectrum expected from this scattering process. Due to Windscheif *et al.*⁶⁹

provided $M^*c_\ell^2$ can be neglected against $\varepsilon = \mathbf{k}^2/2M^*$. According to (11.6.12), the probability that the longitudinal acoustic phonon with energy $\omega_{\text{LA}(\Gamma)}(q) = c_\ell|q|$ is emitted is given⁶⁹ by the shaded lineshape illustrated in Fig. 11.8, with maximum at

$$\omega_{\text{LA}(\Gamma)}_{\text{max}} = 2M^*c_\ell^2[(2\varepsilon/M^*c_\ell^2)^{1/2} - 1] \equiv \omega_{\text{LA}}(\varepsilon). \quad (11.6.14)$$

Further contributions to $\Gamma_t(\varepsilon)$ comes from intervalley scattering $L \rightarrow L'$ by phonons $\omega_s(\mathbf{K}_X)$ of various modes s with $\mathbf{K}'_L - \mathbf{K}_L = \mathbf{K}_X$ at X-points on the Brillouin zone edge at (111) directions.⁶⁶ As a result, the energy Ω_2 of the scattered light is given by

$$\Omega_2 = \Omega_1 - \omega_s(L) - \omega_{s'}(L) - \omega_{s''}(X) - \omega_{\text{LA}}(\varepsilon) \quad (11.6.15)$$

with possible combinations of the last two terms not explicitly written corresponding to possible phonon states ω_n appearing in (11.6.7) and originating from appropriate high-order terms in the expansion (11.6.8).

The observed spectra of Raman shift $\Omega_1 - \Omega_2$ are shown against Ω_1 in Fig. 11.9.⁷⁰ The horizontal lines represent emission of momentum-conserving phonons of modes s and s' and intervalley scattering phonons of mode s'' in various combinations, while the curved lines rising from each horizontal line with threshold at the luminescence line S (with 45° slope) have an energy dependence as given by (11.6.14). The intensities of the Raman lines as functions of Ω_1 , given by the points in Fig. 11.10,⁷¹ can be explained by considering the energy dependence (11.6.13) giving rise to the thresholds as shown by upward arrows, an energy-independent rate Γ_{trap} for being captured at traps and the onset of intervalley scattering giving rise to a sharp drop (note the energy denominator on the r.h.s. of (11.6.11)) at the downward arrows, as shown by the solid lines. The fact that Γ_{trap} is orders of

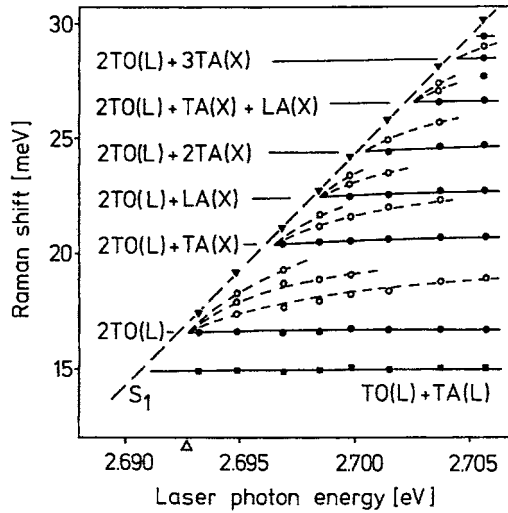


Fig. 11.9 Raman shift of various scattering processes in AgBr involving TA(X) and LA(X) phonons vs excitation photon energy. The diagonal marks the energy position of the luminescence line S_1 . Dispersive LA(Γ) and forbidden TO(L) + TA(L) phonon scattering is also shown. The triangular mark at the abscissa denotes the energy of the TO(L)-assisted exciton absorption edge. Due to Windscheif *et al.*⁷⁰

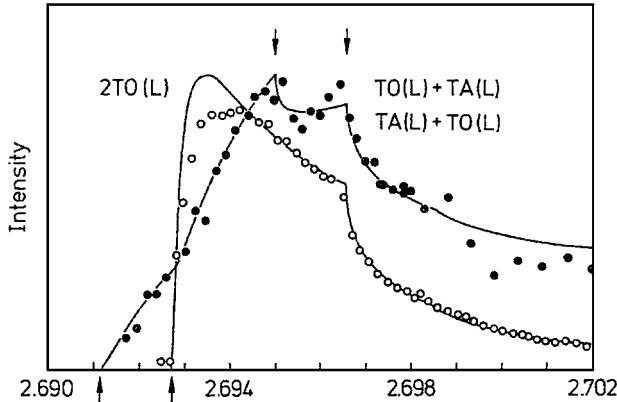


Fig. 11.10 Integrated scattered intensity as a function of excitation photon energy (in eV) for allowed 2TO(L) and partially forbidden TO(L) + TA(L) resonant Raman scattering in AgBr at 1.8 K. The open and full circles are experimental data, the solid lines theoretical fits. The intensities are normalized with the 2TO(L) curve reduced by a factor of 50. Upward arrows mark TA(L)- and TO(L)-assisted absorption thresholds, downward arrows indicate corresponding onsets of intervalley TA(X) scattering. Due to Sliwczuk *et al.*⁷¹

magnitude larger in AgCl than in AgBr⁷² is ascribed to the self-trapping of the exciton which takes place only in the former (see Section 9.5).

Resonance Raman spectroscopy of indirect excitons with its high resolving power was extended to the study of the effects of hydrostatic pressure⁷³ and magnetic

field,⁷⁴ and to the picosecond time-resolved observation⁷⁵ of various relaxation processes including even the dephasing process through quantum beats⁷⁶ (the first observation in solids) between internal levels with different spin states, revealing the fine detail and rich physics of indirect excitons in silver halides.

11.7 Giant two-photon absorption and hyper-Raman scattering associated with an excitonic molecule

Under intense laser light, one can readily obtain a significant density of excitons even in a direct-gap material with a short radiative lifetime. What happens when their density is so high that they have a good chance of coming close to one another? In the analogy of a Wannier exciton X to a hydrogen atom H , one may well imagine the possibility of two excitons forming an *excitonic molecule* (*biexciton*) X_2 similar to a hydrogen molecule H_2 . A great difference of the latter from the former is that a proton is so much heavier (by 1840 times) than an electron that one can safely apply the adiabatic approximation with electrons moving around the protons as fixed point charges, thus forming the covalent bond between two hydrogen atoms. Since the effective masses, m_e and m_h , of the conduction-band electron and the valence-band hole are both of the same order as the electron mass, the situation is less favorable for the formation of a molecule: the zero-point vibrational energy of the heavier particles may exceed the attractive potential mediated by the lighter particles, even if the adiabatic approximation is admitted to be tolerable in such a situation with mass ratio of the order of unity.

In fact, the excitonic molecule was experimentally confirmed in $CuCl$ under intense laser light, by the luminescence emitted when one exciton of the molecule is annihilated leaving a single exciton^{78,79}. Since an excitonic molecule is expected to have an effective mass approximately twice as large as that of a single exciton, m_{ex} , one can write the energy Ω of a photon emitted with conservation of momentum $\hbar\mathbf{K}$ of the excitonic system as

$$\begin{aligned}\Omega &= [E(X_2) + \hbar^2\mathbf{K}^2/4m_{ex}] - [E(X) + \hbar^2\mathbf{K}^2/2m_{ex}] \\ &= E(X) - B(X_2) - \hbar^2\mathbf{K}^2/4m_{ex}\end{aligned}\quad (11.7.1)$$

where

$$B(X_2) \equiv 2E(X) - E(X_2) \quad (11.7.2)$$

is the binding energy of the excitonic molecule. Assuming that the molecules in the initial states follow the Boltzmann distribution, one expects the emission spectra to have an *inverted* Boltzmann distribution (*longer* tail on the *lower-energy* side):⁷⁹

$$I(\Omega) \propto [E(X) - B(X_2) - \Omega]^{1/2} \exp\{-[E(X) - B(X_2) - \Omega]/k_B T\}. \quad (11.7.3)$$

The agreement of the observed lineshape with (11.7.3) (corrected by linewidth) and the consistency of the $B(X_2)$ value (obtained from the observed peak position) with theoretical estimation, together with the observed quadratic dependence of the emission intensity on the exciton density provided convincing evidence for the excitonic molecule in this material.

Akimoto and Hanamura,⁸⁰ and Brinkmann, Rice and Bells⁸¹ carried out a variational calculation of the ground-state energy $E(2e, 2h)$ of the system, consisting of two electrons with spin $1/2$ and isotropic mass m_e and two positive holes with spin $1/2$ and isotropic mass m_h , interacting with each other through Coulomb potentials, with the use of various trial wave functions (reminiscent of hydrogen molecular wave function for certain limiting values of variational parameters). The spin multiplicity of the ground state turns out to consist of the electron singlet and the hole singlet, with no exchange interaction between electrons and holes since they are treated as particles of different species in this model. The obtained binding energy of the excitonic molecule, given by $B(X_2) = 2E(e, h) - E(2e, 2h)$ where $-E(e, h) \equiv Ry_{ex}$ (see (8.2.1)) is the electron-hole binding energy in an exciton, is positive for any value of effective mass ratio $\sigma \equiv m_e/m_h$, namely, an excitonic molecule is always more stable than a distant pair of excitons in materials with such simple direct bandgap structures, of which CuCl is an example. The binding energy of an excitonic molecule, normalized to that of an exciton, namely $B(X_2)/Ry_{ex}$, is a monotonically decreasing function of σ in the interval $[0, 1]$. For the most recent calculation we refer here to the work of Usukura, Suzuki and Varga⁸² who made use of the stochastic variational method with correlated Gaussian basis which is widely used in atomic and molecular physics. According to their result, $B(X_2)/Ry_{ex} = 0.348$ at $\sigma = 0$ (the case of H_2) and $= 0.032$ at $\sigma = 1$ (the case of Ps_2 , the positronium molecule which has not yet been observed). Alongside this, they confirmed two excited states of the excitonic molecule with positive binding energies for any σ .

In materials with more complicated band structures at the direct or indirect band edges, the electronic structure of excitonic molecules are more varied. In materials with indirect band edges, the electron-hole metallic drops are more stable than the gas of excitonic molecules at high e-h concentrations. Referring to a review article⁸³ for more details of excitonic molecules in different materials and many-body effects in high-density exciton systems, we will confine ourselves to the simplest case of CuCl mentioned above, and describe qualitatively the giant cross-sections of two-photon absorption and hyper-Raman scattering of which the excitonic molecule is the final and the intermediate state, respectively.

Let us consider the creation of an excitonic molecule by a single process of two-photon absorption, instead of creating two independent excitons by successive

one-photon absorption. The photon energy Ω' needed for the former, given by

$$2\Omega' = 2(\varepsilon_g - Ry_{\text{ex}}) - B(X_2), \quad (11.7.4)$$

is *slightly* smaller than that for the latter, $\Omega = \varepsilon_g - Ry_{\text{ex}}$, since the binding energy $B(X_2)$ of an excitonic molecule, being only a fraction (0.03–0.35, see above) of the exciton Rydberg energy Ry_{ex} , is much smaller than the energy $\varepsilon_g - Ry_{\text{ex}}$ of an exciton itself. In the expression (11.2.1) with (11.2.2) for the rate of two-photon absorption, the most important intermediate state ℓ is a single exciton with $E_{\ell m} = \varepsilon_g - Ry_{\text{ex}}$, with energy denominator $B(X_2)/2$ being much smaller than Ω' . In addition to this enhancement factor, one has another factor due to the matrix element $P_{f\ell}$ for the second stage in which the second exciton is formed around the first exciton as its bound state: the giant oscillator strength for the creation of a bound exciton pointed out by Rashba and Gurevishvili (see Ref. [21] of chapter 8, in particular eq. (8.8.12)). Altogether, one has an enormous enhancement factor proportional to $[B(X_2)/\varepsilon_g]^{-7/4}$ due to the smallness of binding energy of the excitonic molecule, as pointed out by Hanamura.⁸⁴

It is to be noted that the second enhancement factor in the giant two-photon absorption, namely the giant oscillator strength for the bound exciton, plays a role also in the *optical conversion* of an exciton to an excitonic molecule,⁸⁵ the inverse process of the emission of an excitonic molecule mentioned before. For example, the optical conversion of excitons with density as low as 10^{17} cm^{-3} in CdS takes place at a rate comparable to the intrinsic absorption of the host crystal with 10^{22} cm^{-3} unit cells.

The first experiment of two-photon absorption for an excitonic molecule with giant cross-section was done by Gale and Mysyrowicz in CuCl,⁸⁶ while Nozue *et al.*⁸⁷ confirmed the excitonic molecule in CdS by giant two-photon absorption.

Giant two-photon absorption (GTA) opened a way to the experimental and theoretical studies of a variety of high-order optical processes. Among them we will describe here the hyper-Raman scattering and luminescence which are emitted from the excitonic molecule as intermediate or initial state and an exciton as the final state. Here again, we encounter the problem of the coexistence of light scattering and luminescence which was already studied in Section 11.6.

Nagasawa *et al.*⁸⁸ confirmed the efficient generation of excitonic molecules by GTA in CuCl and CuBr through the investigation of excitation spectra for the emission band therefrom. The emission band consists of M_L - and M_T -bands (with the inverted Boltzmann distributions) in the case of CuCl, corresponding to the transition from the excitonic molecule (EM) to a longitudinal exciton (X_L) and a transverse exciton (X_T), respectively. While the threshold position (see (11.7.3)) of the former band was found to be exactly equal to the energy difference $E(X_2) - E(X_L)$;

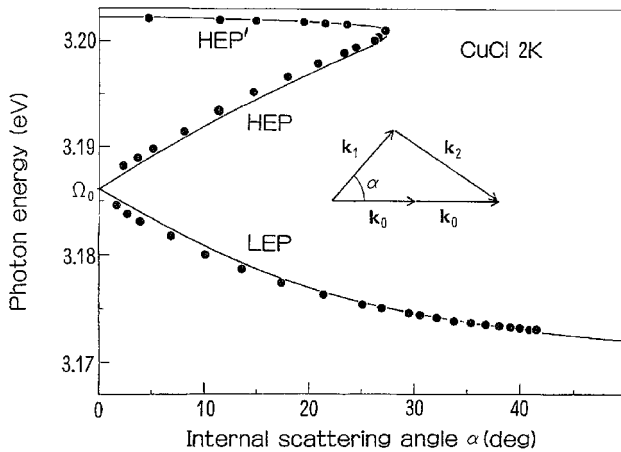


Fig. 11.11 Energies of hyper-Raman lines associated with excitonic molecule as intermediate state in CuCl at 2 K vs internal scattering angle. Due to Ueta *et al.*⁸⁹ Theory (solid lines) is due to Inoue and Hanamura.⁹⁰

the latter appeared at a somewhat lower energy than $E(X_2) - E(X_T)$ which was ascribed to the polariton effect of X_T . With the use of laser light with a much narrower spectral width, Ueta *et al.*⁸⁹ found a number of Raman lines: M_L^R , M_T^R , LEPs, HEPs (see Fig. 11.11), UP(II) and so on, depending upon the experimental configurations: directions of two incident beams and of observation, in addition to the emission bands M_L and M_T mentioned above. The observed dependence of the energies of these Raman lines on the configurations as shown in Fig. 11.11 was explained by the theory of Inoue and Hanamura⁹⁰ based on the use of energy and momentum conservations in the four-photon process with two incident beams $\Omega_1, \mathbf{k}_1, \mathbf{k}_1'$ and two emitted polaritons $\omega(\mathbf{K}), \omega'(\mathbf{K}')$:

$$2\Omega_1 = \omega(\mathbf{K}) + \omega'(\mathbf{K}'), \quad \mathbf{k}_1 + \mathbf{k}_1' = \mathbf{K} + \mathbf{K}', \quad (11.7.5)$$

where the polariton satisfies the dispersion relation (see (8.6.8)):

$$c_0^2 K^2 / \omega^2 = (\epsilon' / \epsilon_0) (\omega_L^2 - \omega^2) / (\omega_T^2 - \omega^2). \quad (11.7.6)$$

(Note that $\omega_L \equiv E(X_L)$, $\omega_T \equiv E(X_T)$.)

11.8 Seeking materials with large nonlinear susceptibilities

The two factors for the enhancement of the cross-section of higher-order optical processes such as was seen in the previous sections, namely (i) a small energy denominator due to the near resonance of the intermediate state, and (ii) a large transition dipole matrix element, provide a key strategy in the development of new

materials with large nonlinear susceptibilities, which are expected to play important roles in modern optical technology.

Insulators with quasi-one-dimensional structures are especially suited for this purpose since they satisfy both criteria. The low dimensionality has the effect of enhancing the long-range Coulomb interaction in such a way that the oscillator strength for the interband transition is strongly concentrated at the lowest exciton, as is seen from the fact that the binding energies of a Wannier exciton in the effective mass approximation in 3-, 2-⁹¹ and 1-⁹² dimensions are in the ratio 1 : 4 : ∞. The divergence in the one-dimensional system is to be removed by taking into account the discrete lattice structures of realistic materials, as has been done in conjugated polymers (such as polydiacetylene)⁹³ and polysilanes.⁹⁴ The interband oscillator strength was found in fact to be strongly concentrated at the lowest exciton, and the calculated nonlinear susceptibilities are in reasonable agreement with observations.

Enormous values of the third-order nonlinear susceptibility $\chi^{(3)}$ have recently been reported by Kishida *et al.*⁹⁵ in one-dimensional Mott insulators with strong electron correlation which are ascribed to the second enhancement factor characteristic of these materials. The dominant term of the nonlinear susceptibility concerned, of which the imaginary part represents the d.c. Kerr effect, is given, according to their three-level model (0 for the ground state, 1 and 2 for the excited states to which the transitions from 0 are dipole-allowed and -forbidden, respectively) by

$$\begin{aligned} \chi^{(3)}(-\omega; 0, 0, \omega) \propto & \langle 0|x|1\rangle\langle 1|x|2\rangle\langle 2|x|1\rangle\langle 1|x|0\rangle \\ & \times [(\omega_1 - \omega - i\gamma_1)(\omega_2 - \omega - i\gamma_2)(\omega_1 - \omega - i\gamma_1)]^{-1} \end{aligned} \quad (11.8.1)$$

where x denotes the dipole moment along the one-dimensional axis. By comparing the susceptibilities of the Mott insulators with those of other types of one-dimensional nonlinear materials, the greater values of $\chi^{(3)}$ are unambiguously ascribed to the greater values, in the former, of the transition dipole moment $\langle 2|x|1\rangle$. In the excited states $|1\rangle$ and $|2\rangle$, an electron is transferred from the occupied ligand state to the unoccupied upper Hubbard band states of the transition metal, and $\langle 2|x|1\rangle$ is a measure of the spatial extension of the electron-hole wave function in these states. The large values of $\chi^{(3)}$ are consistent with the large transfer energy resulting from strong hybridization between the transition-metal orbital and the ligand orbitals. An advantage of these Mott insulators with strong electron correlation is, according to the authors, that they can maintain a large optical gap in spite of their large transfer energy.

12

Inner-shell excitation

12.1 Inner-shell electrons and synchrotron radiation

We have so far been concerned with the electrons in the *outermost shell* of atoms which play the main roles in the formation of interatomic bonds resulting in the aggregation of atoms into molecules or condensed matter. Material properties sensitive to external perturbations are governed by these electrons.

In contrast to these *valence* electrons (in insulators) and *conduction* electrons (in metals and semiconductors), the *inner-shell* electrons (also called *core* electrons) with which we shall be concerned in this chapter are more tightly bound around each nucleus and can be assumed to a good approximation to be localized within a particular atom. They are much less influenced by other atoms except through the chemical shift due to surrounding atoms or ions which is more or less common to all inner shells of the same atom (such as the Madelung potential). The inner-shell electrons have been left out of the main stream of molecular and condensed matter spectroscopy, simply because of the rather limited light sources in the regions of energy (vacuum ultraviolet (VUV) to X-ray) needed for the excitation of such deep-lying electrons.

The situation changed completely when *synchrotron radiation* (SR) became accessible to spectroscopic and structural researchers. It is a sort of *Bremsstrahlung*, the electromagnetic radiation emitted by electrons with high velocities in the relativistic regime which are constantly subjected to transverse *acceleration* (due to the Lorentz force generated by the bending magnets) on their circular orbit in the synchrotron. The primary purpose of this machine was to accelerate charged particles for the study of high-energy physics, and SR had been the main mechanism of energy loss which needed significant microwave power to compensate it, but this undesirable byproduct soon turned out to be not only a useful but an indispensable tool for the spectroscopic and structural studies of matter.

Due to the relativistic effect, studied in detail by Ivanenko and Pomerachuk¹ and Schwinger,² the emitted radiation is concentrated mostly in the forward direction of the instantaneous velocity of the electron, and is strongly polarized in the direction of the instantaneous acceleration (towards the center of the circular orbit).³ The spectral distribution of the emitted radiation is not confined to the fundamental frequency of the revolution (\sim MHz) but includes its harmonics (each being smeared out) up to very high order due to the almost point-like charge electrons of relativistic velocities, providing a light source with a continuous spectrum extending from infrared through visible and UV up to X-ray regions which is extremely useful for the spectroscopic research of inner-shell electrons whose binding energies are widely distributed from several to thousands of electron volts depending on the principal quantum number n of the shell as well as on the atomic number Z . The total power of the radiation emitted by an electron is proportional to E^4/R^2 while its spectral peak is at a wavelength proportional to R/E^3 where E is the stationary energy of the accelerated electron and R the radius of its circular orbit. As for the basic principles of this device and its applications to inner-shell spectroscopy see, for instance, Ref. [4].

Since the electronic states of inner shells are well known in atomic physics and are relatively little influenced by surrounding atoms, as mentioned above, the main interest of inner-shell spectroscopy in molecules and solids consists in studying (I) the final electronic states of the optical transition which depend much more on the nature of interatomic aggregation, and (II) a variety of electronic and atomic processes inclusive of (II α) *simultaneous* excitation of various modes of elementary excitation in aggregates as reflected in the spectral width and sidebands (satellites) and (II β) *subsequent* atomic displacements such as lattice relaxation, defect formation and chemical reaction. Since the energies of different inner shells (named K, L, M, N, O, P, Q according as $n = 1, 2, 3, \dots, 7$) and subshells ($\ell = 0, 1, 2, 3, \dots$ denoted by s, p, d, f, \dots) of different atoms are widely separated from each other, the thresholds of the corresponding absorption spectra are clearly distinguishable. The *wide* spectral range of SR facilitates a *comparative* study of the optical spectra due to *different shells and subshells of different atoms*, providing useful *complementary* knowledge on the electronic structure of the final states (item (I) mentioned above) inclusive of their *symmetry* around the *particular atom* excited (*selection of atom and symmetry*). As an example of such a comparison, one can think of the different lineshapes at the allowed and forbidden band edges (described in Section 8.4(a) and (b)) depending on the initial subshell. For the same reason, one can realize *selective* desorption of an atom of a *particular species* out of a compound material by *selective excitation* of the inner (sub)shell of an atom of the *appropriate species* or *site* (the latter only in the case of *inequivalent* sites with different chemical shifts), which will also shed light on the mechanism by which the excited atom

transfers its energy to cause that nearby atom to be desorbed (item (II β) mentioned above).

With SR spectroscopy, one can also study the *entire* range of the absorption spectrum of a particular inner shell, which extends from its threshold over the range of energy of the same order as the binding energy of that shell, and which consists of *hierarchical* structures reflecting the short- and long-range arrays of atoms in the aggregate as well as the nodal structure of the inner-shell wave function – an interplay of *local and band* structures, so to speak – as will be described in the next section.

12.2 Hierarchical structures in absorption spectra of inner shells

In order to obtain a basic understanding of inner-shell absorption spectra in atoms, molecules and solids in a unified way, we will study here the *redistribution* of oscillator strengths in an *aggregate* of atoms^{5,6} referred to those of the *isolated* atoms which were studied in detail by Cooper and collaborators.⁷⁻⁹ Denoting the atomic, molecular or valence orbitals of *occupied* states by ϕ_μ , we want to see how to find the orbitals of *unoccupied* states ϕ_α . If the aggregate is electronically *insulating* with a *finite* energy gap between the highest occupied states and the lowest unoccupied states, it will be obvious to us, before solving for ϕ_α , what ϕ_μ we should take. In the first approximation, the ϕ_μ s consist of all atomic orbitals $\phi_{\mu i}(\mathbf{r}) \equiv \phi_\mu(\mathbf{r} - \mathbf{R}_i)$ of the occupied states with quantum number $\mu (\equiv (n, \ell, m, s))$ of atoms (or ions in the case of ionic compounds) at \mathbf{R}_i and all occupied bonding orbitals (in the case of covalent molecules or crystals). One can write the Schrödinger equation for the unoccupied states in the aggregate as

$$[-(\hbar^2/2m)\nabla^2 + v(\mathbf{r})]\phi_v = \varepsilon_v \phi_v, \quad (12.2.1)$$

where $v(\mathbf{r})$ is approximately given by the superposition of potentials $V_i(\mathbf{r} - \mathbf{R}_i)$ of the constituent atoms (or ions) i at \mathbf{R}_i with their electrons. The suffix i to V is to distinguish the *species* of atom as is necessary in the case of *compound* aggregates. Any unoccupied state ϕ_v must be orthogonal to the occupied states ϕ_μ and hence can be written in the form

$$\phi_v(\mathbf{r}) = \chi_v(\mathbf{r}) - \sum'_\mu (\phi_\mu, \chi_v) \phi_\mu(\mathbf{r}) \quad (12.2.2)$$

where \sum' extends over all occupied states. Insert (12.2.2) into (12.2.1) and impose a condition that $\chi_v(\mathbf{r})$ be as smooth as possible, following the prescription of Cohen and Heine¹⁰ which is appropriate for such a situation. Then one obtains the so-called *pseudo-potential* v_p , to be used in the effective Hamiltonian for χ_v , as the following

integral operator

$$v_p(\mathbf{r}, \mathbf{r}') = v(\mathbf{r}') \left[\delta(\mathbf{r} - \mathbf{r}') - \sum_{\mu}^i \phi_{\mu}^*(\mathbf{r}') \phi_{\mu}(\mathbf{r}) \right], \quad (12.2.3)$$

where the second term in the square bracket arises from the orthogonalization. As seen from the fact that v_p given by (12.2.3) would vanish if the ϕ_{μ} s formed a complete set, v_p is supposed to be much smaller in magnitude than the true potential v . This is the *cancellation of potential*, well known in solid state physics.¹⁰ The effective Schrödinger equation for χ_v is thus expressed as

$$[-(\hbar^2/2m) + v_p]\chi_v = \varepsilon_v \chi_v. \quad (12.2.4)$$

The orthogonalization gives another important effect, the *cancellation of oscillator strength*⁵ in the absorption spectrum, as will be seen below.

(i) *Cancellation of oscillator strength – giant and subgiant bands*

In so far as we consider the gross structure of the absorption spectra in the energy region of several tens of eV above the threshold, one can simply replace v_p by its average v_0 (constant) over the entire space to obtain the solution of (12.2.4) as plane waves $\chi_{\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r})$ with energies $\varepsilon_0(\mathbf{k}) = v_0 + \hbar^2 k^2/2m$. To make the mathematical analysis more transparent, we will confine ourselves, within this subsection, to the aggregate of atoms with closed-shell structures such as rare-gas solids and ionic crystals, and neglect the interatomic overlap of occupied-shell orbitals. The oscillator strength of an electron in the $(n\ell)$ subshell of an atom i to the final state continuum per unit energy is given, according to the definition (3.6.6), by

$$f(n\ell i \rightarrow k) = (\sqrt{2}m^{5/2}/3\pi^2\hbar^5)\hbar\omega(\hbar\omega - \hbar\omega_i)^{1/2} \\ \times \sum_{\ell'=\ell\pm 1} [(\ell + \ell' + 1)/2(2\ell + 1)][R(n\ell i, k\ell')]^2, \quad (12.2.5)$$

$$R(n\ell i, k\ell') \equiv \int_0^{\infty} r u_{n\ell i}(r) v_{n\ell'}(r) dr, \quad (12.2.6)$$

$$v_{k\ell'}(r) \equiv w_{k\ell'}(r) - \sum_n^i \langle u_n^{\ell'} | i | w_{k\ell'} \rangle u_n^{\ell'}(r), \quad (12.2.7)$$

$$w_{k\ell'}(r) = (4\pi)^{1/2} i^{\ell'} j_{\ell'}(kr). \quad (12.2.8)$$

Here $u_{n\ell}$ and $w_{k\ell}$ are r times the atomic (ionic) radial wave functions of the occupied n -shell and the ionized state k , respectively, and $j_{\ell}(kr)$ is the Bessel function. The function $u_{n\ell}$ satisfies the Schrödinger equation

$$\{-(\hbar^2/2m)(d/dr)^2 + [V_i(r) + \hbar^2\ell(\ell+1)/2mr^2 + \zeta_i]\}u_{n\ell i} = E u_{n\ell i} \quad (12.2.9)$$

where the second term in $[\dots]$ is the *centrifugal potential* originating from the angular part of the kinetic energy operator, $-(\hbar^2/2m)\nabla^2$. We have taken the potential $\zeta(\mathbf{r})$ due to the surrounding atoms i' at $\mathbf{R}_{i'} \neq \mathbf{R}_i$, which is given by

$$\zeta(\mathbf{r}) = \sum_{i'} V_{i'}(\mathbf{r} - \mathbf{R}_{i'}), \quad (12.2.10)$$

to be a constant ζ_i within the atom i , so that we can replace $u_{n\ell i}$ by that of an isolated atom with energy $\varepsilon_{n\ell i}$, and E by $\varepsilon_{n\ell i} + \zeta_i$. The constant ζ_i is called chemical shift in the case of compound aggregates. As an example, the Madelung potential in ionic crystals amounts to the order of eV. In the similar equation for the extended state $w_{k\ell}$, we take, instead of ζ_i , the average v_0 of v_p over the entire space as already mentioned. (Note that the contributions to (12.2.10) of the orthogonalization terms in (12.2.3) from other atoms have been neglected within the atom of our concern, as they can be.) Therefore, the threshold of the photon energy $\hbar\omega$ for ionization of atom i , to be used in the formula (12.2.5), is given by

$$\hbar\omega_t' = v_0 - (\varepsilon_{n\ell i} + \zeta_i), \quad (12.2.11)$$

although a more exact threshold, $\hbar\omega_t$, with the effect of $v_p - v_0$ taken into account, will be given in (ii) below.

The oscillator strengths for the spherical wave components of the unorthogonalized plane waves, i.e. $f(n\ell i \rightarrow k, \ell' = \ell \pm 1)$ obtained by putting $v_{k\ell'} = w_{k\ell'}$, are shown schematically by the thin and thick solid curves in Fig. 12.1. The maximum

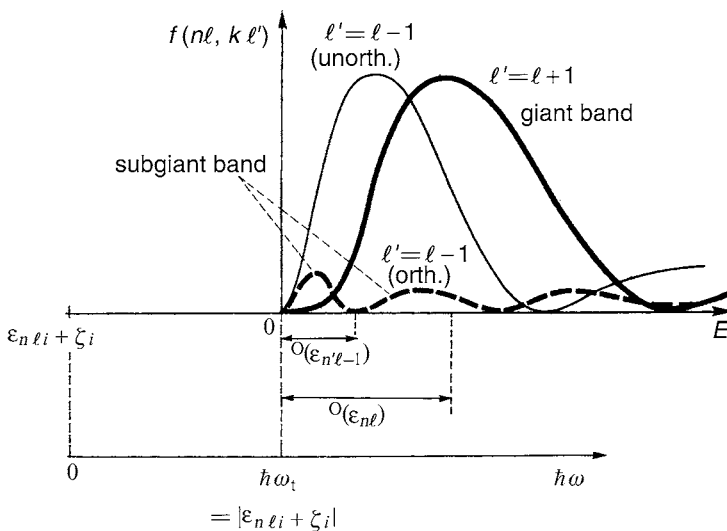


Fig. 12.1 Schematic behavior of oscillator strengths from the $(n\ell)$ -shell to $\ell' = \ell \pm 1$ components of plane-wave states. The broken curve is for the orthogonalized plane-wave state.

appears at higher energy for $\ell' = \ell + 1$ than for $\ell' = \ell - 1$, corresponding to the larger centrifugal potential $\hbar^2 \ell'(\ell' + 1)/2mr^2$ to be added to the attractive potential from the atomic core in the Schrödinger equation for the radial wave function $w_{k\ell'}(r)$. The average extension of the spectra beyond $\hbar\omega_t$ is of the same order of magnitude as the binding energy $|\varepsilon_{n\ell}|$ of the initial state since $R(n\ell, k\ell')$ is appreciable only for those $v_{k\ell'}(r)$ with wavelength $2\pi/k$ being at least of the same order of magnitude as the orbital radius of the initial state. Furthermore, if $n > \ell + 1$, nodes appear in the oscillator strength versus energy curve. These features are well known in atoms.

The same features can be described in a more transparent way by resorting to the pseudo-potential theory, as will be shown below. In the first place, the effect of orthogonalization of the final-state wave function, represented by the second term of (12.2.2), gives rise to cancellation of the oscillator strength,^{5,6} a counterpart of the cancellation of the potential well known in the pseudo-potential theory.¹⁰ In general, the number of filled shells (n', ℓ') to which the final state ($k\ell'$) has to be orthogonalized, and hence the degree of cancellation of the oscillator strength, is greater for $\ell' = \ell - 1$ than for $\ell' = \ell + 1$. For instance, in the case of 4d-shell photoabsorption of I, Xe and Cs, one has four filled p-shells (2p–5p), whereas the lowest 4f-shell is still empty. Orthogonalization of the wave function to these shells reduces the oscillator strength $f(4d \rightarrow kp)$ by one or two orders of magnitude, while there is no filled f-shell with which the (kf) states must be orthogonalized. These results based on orthogonalization provide an alternative explanation to that found in the case of an isolated atom, in a form applicable also to an aggregate of atoms. Furthermore, the subgiant band is found to oscillate, as shown by the thick broken curve in Fig. 12.1, because the orthogonalized spherical wave state (12.2.7) has nodes near the atomic core. The separation of the first node of the subgiant band from $\hbar\omega_t$ is of the order of the binding energy $|\varepsilon_{n', \ell-1}|$ of the *outermost occupied* electron with quantum number $\ell - 1$. In less ionic and more covalent crystals, this part of the subgiant band may be subject to some change due to the existence of valence bonds as the outermost *partially* occupied $\ell - 1$ orbitals which has not been considered in the above simplified description.

The characteristics as depicted in Fig. 12.1 can be seen most clearly in the spectra of (nd) core excitation in closed-shell insulators such as rare-gas solids and alkali halides^{11–13} as shown in Fig. 12.2 for the Rb-3d core excitation in RbCl and the I-4d core excitation in KI. One finds a broad and intense giant band, below which a weak subgiant band with width of the same order as the binding energy of the highest occupied p-shell of the atom, in accordance with the above-mentioned result of the pseudo-potential theory.

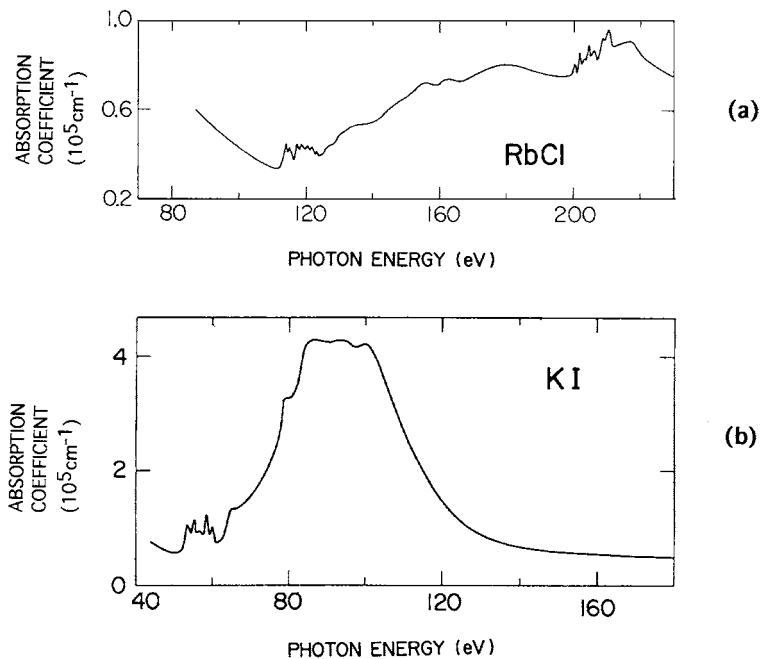


Fig. 12.2 Absorption spectra of (a) Rb-3d core (with Cl-2p to the right) in RbCl and (b) I-4d core in KI. Due to Brown *et al.*¹³

(ii) Pseudo-potential and the energy band

For an isolated atom, where $v(\mathbf{r}) = V(\mathbf{r})$, the pseudo-potential for the final state with ℓ' is given by

$$v_p^{(\ell')}(\mathbf{r}, \mathbf{r}') = V(\mathbf{r}') \left[\delta(\mathbf{r} - \mathbf{r}') - \sum_{n', m'} \phi_{n' \ell' m'}^*(\mathbf{r}') \phi_{n' \ell' m'}(\mathbf{r}) \right]. \quad (12.2.12)$$

As we increase the atomic number Z , $v_p^{(\ell')}$ as well as V become deeper, until a new shell (n'' , ℓ') begins to be occupied. Consequently, the peak of the giant band is expected to shift towards the threshold, since the effect of the centrifugal repulsion is reduced due to $v_p^{(\ell')}$. The $4d \rightarrow \varepsilon f$ giant band for the sequence I^- , Xe, Cs^+ (with $Z = 53, 54, 55$, respectively) shows this tendency¹¹⁻¹³ which is a typical example of the *resonance near threshold* found in an atom.⁷⁻⁹ When a new subshell ℓ' is filled up, we have to include it under the summation in (12.2.12), so that $v_p^{(\ell')}$ becomes shallow again.

In molecules and solids, there appears, in addition to the shift of the giant band mentioned above (intra-atomic effect), another effect (interatomic) which comes from the pseudo-potential (12.2.3) instead of (12.2.12). The solutions ϕ_v of (12.2.1), obtained by putting the solution χ_v of the effective equation (12.2.4) into (12.2.2), are now expressed as molecular orbitals, or Bloch orbitals $\phi_{vk}(\mathbf{r})$ with associated

energy band $\varepsilon_v(\mathbf{k}) = \varepsilon_v + \hbar^2 \mathbf{k}(m_c^{-1})\mathbf{k}/2 + \dots$ which is to replace $\varepsilon_0(\mathbf{k})$ mentioned in (i) above (m_c is the effective mass tensor). These orbitals are in fact orthogonalized to those of the occupied states due to (12.2.2). For example, we would obtain an antibonding state ϕ_v which is orthogonal to the corresponding bonding state ϕ_μ for a covalent bond. This gives rise to fine structures, characteristic of the respective *aggregates*, superimposed on the subgiant bands of atomic origin. It is in this part of the spectra that comparison with band theory has been successful.^{14,15} (See also p. 57 of Ref. [16]). The energy scale of these fine structures is of the order of eV, in contrast to those of the giant (several tens of eV) and subgiant (~ 10 eV) bands. The exact threshold $\hbar\omega_t$ is given by

$$\hbar\omega_t = \varepsilon_v - (\varepsilon_{n\ell i} + \zeta_i) \quad (12.2.13)$$

with the use of the band-bottom energy ε_v in place of v_0 in (12.2.11).

(ii) *The electronic-polaron effect*

The above-mentioned one-electron approximation should further be corrected for many-body effects, especially the effect of electronic polarization around the excited electron in the conduction band and the positive hole left in the inner shell ($n\ell i$). Namely, their electronic-polaron self-energies (see Section 9.3), Δ_v and $\Delta_{n\ell i}$ both of which are negative and on the order of eV, should be added to the bare energies ε_v and $-\varepsilon_{n\ell i}$, respectively, so that the renormalized energies, $\bar{\varepsilon}_v = \varepsilon_v + \Delta_v$ and $\bar{\varepsilon}_{n\ell i} = \varepsilon_{n\ell i} - \Delta_{n\ell i}$, are to be used in (12.2.13), resulting in a decrease in the threshold energy of the order of a few eVs. It is to be noted, incidentally, that the self-energy of the inner-shell hole belongs definitely to the classical case (c) and is given by (9.3.23) since the corresponding band has negligible dispersion.

(iii) *Effect of electron-hole Coulomb attraction*

If the attractive Coulomb potential v_c due to the positive hole, so far neglected in the one-electron approximation, is taken into account, an infinite number of bound states appear below the ionization threshold; they are Rydberg states in atoms and molecules, and excitons in solids which we have described in detail in Chapter 8. Usually, v_c is smaller in solids because of dielectric screening. The oscillator strengths of Rydberg states are significantly reduced in excitons, approximately by $(m_c/m)(\varepsilon_{\text{eff}}/\varepsilon_0)^{-3}$ where ε_{eff} is an effective dielectric constant somewhere between ε_e and ε_0 , the remainder being removed into the continuum of band-to-band transition. The effective mass m_c of the conduction band, instead of the reduced mass $\mu = (m_c^{-1} + m_v^{-1})^{-1}$, is to be used since the positive hole in the inner shell has enormous effective mass (negligible interatomic transfer energy).

The fact that v_c is small especially in solids warrants our successive approximation above – (i) orthogonalized plane wave, (ii) pseudo-potential with (ii') electronic-polaron effect, and then (iii) the Coulomb effect.

(0) *Extended X-ray absorption fine structure (EXAFS)*

In the inner-shell absorption spectra of molecules and solids a series of small-amplitude oscillations with period of the order of 50 eV is observed up to 1000 eV above the threshold. This oscillation arises from the scattering of outgoing waves by the pseudo-potential v_p , and has the same origin as that of the Bloch bands in crystals when described by the nearly free electron model.

With large values of wave vector \mathbf{k} corresponding to such high energies, the scattered waves of the electron by atoms *distant* from the site of its inner-shell excitation cannot effectively interfere with the primary outgoing wave to cause spectral structures because of their smeared-out phases. Namely, more or less independent vibrations of different atoms cause irregular change in the relative phases of the scattered waves. More important are scatterings by nearby atoms which give rise to the above-mentioned spectral structure, as will now be described very briefly.¹⁷

The outgoing wave of an electron excited from the inner shell of the atom at $\mathbf{R} = 0$ will be scattered by a nearby atom at \mathbf{R}_j with scattering amplitude $S(\theta)$, so that the back-scattered ($\theta = \pi$) wave will give, at $\mathbf{R} = 0$, an additional contribution to the oscillator strength $f_0(\omega)$ of the primary wave given by (12.2.5):

$$f(\omega) = f_0(\omega)[1 + g(k)], \quad (12.2.14)$$

$$g(k) = |S(\pi)| \sin(2kR_j + \phi) / (kR_j^2), \quad (12.2.15)$$

$$S(\pi) = |S(\pi)| \exp(i\phi). \quad (12.2.16)$$

Note that $2R_j$ and ϕ represent the distance of the return trip to the atom j and the phase shift of the scattering thereby, respectively. Thus the period of EXAFS is given by $\Delta k = \pi/R_j$ (the same as the first Brillouin zone boundary well known in the band theory), which corresponds to $\Delta(\hbar\omega) = (\pi\hbar/R_j)[2(\hbar\omega - \hbar\omega_l)/m]^{1/2}$ since $(\hbar\omega - \hbar\omega_l) = \hbar^2 k^2/2m$. Assuming $R_j \sim 2.5 \text{ \AA}$ (nearest neighbor distance for Cu), for instance, we obtain typical values of the period $\Delta(\hbar\omega) \sim 50 \text{ eV}$ and 150 eV for $(\hbar\omega - \hbar\omega_l) \sim 100 \text{ eV}$ and 1000 eV, respectively.

Summing up the contributions of all constituent atoms, we finally obtain

$$g(k) = \sum_j N_j / (kR_j^2) |S_j(\pi)| \sin(2kR_j + 2\delta_l' + \phi_j) \exp(-2\sigma_j^2 k^2) \exp(-\gamma R_j) \quad (12.2.17)$$

where the first exponential is the Debye–Waller factor representing the above-mentioned dephasing effect of the lattice vibrations, the second exponential is to

account for the effect of inelastic scattering by the damping factor, and δ_1^1 represents the effect of the potential of the excited atom itself due to the core hole left behind.

EXAFS, with its period reflecting the distance R_j , provides us with knowledge of the atomic arrangement in materials, as does low-energy electron diffraction (LEED) or conventional X-ray diffraction. However, EXAFS has many advantages over the last two: it is particularly sensitive to the *local environment* of atoms (short-, rather than long range) and the surroundings of *each species* of atoms can be separately studied by tuning the corresponding absorption edge. Furthermore, unlike LEED, complicated multiple scatterings average out and give no serious effect on the lineshape except for some special cases. In view of these facts, as well as the development of the SR source, there has recently been much interest in theoretical and experimental studies of EXAFS, and its application to the structural investigation of various materials has been proposed and performed. For instance, EXAFS is expected to be especially useful in the study of non-crystalline systems^{18,19} and complex systems containing various chemical elements, such as biomolecules and catalysts.²⁰

12.3 Auger process and its competition with the radiative process

Inner-shell excitation with high energy can invoke a variety of secondary processes by which a part of the excitation energy is converted into other channels of excitation. Among them, the electronic processes take place in general more rapidly than the atomic processes such as chemical reactions and lattice relaxations because the atomic mass is much heavier than the electronic mass. Since the positive hole left behind in a *deep* inner shell (also called a core hole) is potentially much more unstable (energetically higher and spatially more confined) than the excited electron in a *shallow* level, the most efficient channels of conversion are associated with the annihilation of the core hole by an electron in outer (less-inner) shells. One is the *Auger process*²¹ in which another electron-hole pair is excited through the Coulomb interaction as shown schematically in Fig. 12.3, another is the radiative process in which the energy is converted into a photon. These two processes contribute additively to the lifetime broadening of the core hole which is much greater than that of outer-shell excitation, and manifest themselves in the absorption and photoelectron spectra of the inner shell. The long-range Coulomb interaction $v(\mathbf{r} - \mathbf{r}')$ allows the Auger transfer of the excitation energy to an electron-hole pair at other atoms (though with smaller rates) thus contributing to the spatial diffusion of the excitation energy, while radiative annihilation is confined to electrons of the same atom (or at most the nearest neighbors as will be seen in an example mentioned later in this section) due to the short-range electron-photon interaction (see Section 3.3).

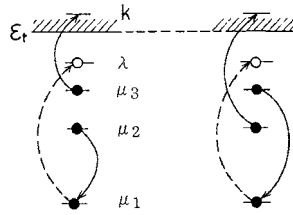


Fig. 12.3 Schematic diagram for the Auger processes (solid arrows). Radiative transitions are shown by dashed arrows. Solid circles represent the core electrons occupied in the ground state.⁶

The matrix element for the Auger process corresponding to Fig. 12.3 consists of two terms due to the indistinguishability of electrons:

$$V_{k\lambda} = \iint d\mathbf{r} d\mathbf{r}' \phi_{\mu_1}^*(\mathbf{r}) \phi_k^*(\mathbf{r}') v(\mathbf{r} - \mathbf{r}') \phi_{\mu_3}(\mathbf{r}') \phi_{\mu_2}(\mathbf{r}) - \iint d\mathbf{r} d\mathbf{r}' \phi_{\mu_1}^*(\mathbf{r}) \phi_k^*(\mathbf{r}') v_c(\mathbf{r} - \mathbf{r}') \phi_{\mu_2}(\mathbf{r}') \phi_{\mu_3}(\mathbf{r}). \quad (12.3.1)$$

If the orbital μ_3 or μ_2 belongs to an atom different from that of the core state μ_1 , $V_{k\lambda}$ represents the interatomic transfer of the excitation energy.

The optical absorption spectrum through the dipole matrix element $M_{\lambda\mu_1} \equiv M$ is expressed, apart from unimportant factors, as

$$I(\omega) = \pi^{-1} \text{Im}[M^\dagger G_{\lambda\lambda} M] = |M|^2 \pi^{-1} \Gamma(\omega) [\{\hbar\omega - E_\lambda - \Delta(\omega)\}^2 + \Gamma(\omega)^2]^{-1} \quad (12.3.2)$$

where

$$G = [z - (H_0 + V)]^{-1}, \quad (z \equiv \hbar\omega - i\eta, \eta \rightarrow +0, \\ H_0 : \text{unperturbed Hamiltonian}) \quad (12.3.3)$$

$$\Delta(\omega) = P \sum_k |V_{k\lambda}|^2 (\hbar\omega - E_k)^{-1}, \quad (12.3.4)$$

$$\Gamma(\omega) = \pi \sum_k |V_{k\lambda}|^2 \delta(\hbar\omega - E_k), \quad (12.3.5)$$

$$E_\lambda = \varepsilon_\lambda - \varepsilon_{\mu_1}, \quad E_k = \varepsilon_k + \varepsilon_\lambda + -\varepsilon_{\mu_2} - \varepsilon_{\mu_3}. \quad (12.3.6)$$

When the ω dependence of $\Delta(\omega)$ and $\Gamma(\omega)$ is disregarded, $I(\omega)$ is of Lorentzian lineshape.

In the above calculation, the level width 2Γ is ascribed to the lifetime of the hole in the deep inner shell μ_1 , while the possible transitions of the electron in shallow level λ has been ignored since its contribution is expected to be small. As

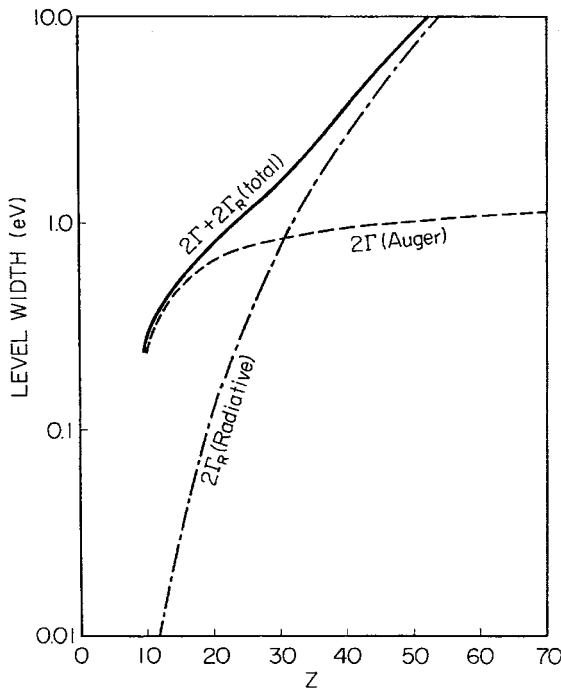


Fig. 12.4 K-level width due to the Auger and radiative processes as a function of atomic number Z .⁶

an example, 2Γ of the K-shell hole in an isolated atom, which is obtained by more elaborate calculation,²² is shown in Fig. 12.4 as a function of atomic number Z . The Auger width is found to increase monotonically with increase in Z .

As mentioned before, another contribution to the level width 2Γ of the core hole comes from its radiative annihilation by an electron occupying the level $\varepsilon_{\mu'}$ higher than the inner shell ε_{μ_1} concerned. One has only to sum over all possible μ' the contribution $\hbar A_{\mu'}$ where the rate A is given by (3.6.7) with $\hbar\omega_{\mu'_1} = \varepsilon_{\mu'} - \varepsilon_{\mu_1}$ and the corresponding oscillator strength $f_{1\mu'}$. In Fig. 12.4, the radiative width $2\Gamma_R$ of the K-shell in an isolated atom is also plotted. The width $2\Gamma_R$ is found to be approximately proportional to Z^4 , which is reasonable since the binding energy of the K-shell is proportional to Z^2 according to the hydrogen model and (3.6.7) contains the factor $(\hbar\omega_{\mu'_1})^2$. In contrast, the Auger width changes with Z more gradually, so that the two curves in Fig. 12.4 cross each other in the neighborhood of $Z \sim 31$.

For core states shallower than the K-shell the contribution of the Auger process to the total width becomes more dominant. In fact, in isolated atoms as well as monatomic crystals, the energy gap between successive inner shells is a rapidly

decreasing function of the principal quantum number n , so that the radiative annihilation of a core hole is usually governed by the Auger processes associated with electrons in shallower shells (see Fig. 12.3). This simple rule does not apply to compound materials since the energy gap between the core states of atoms of different species is a completely independent parameter. Some ionic crystals have been found in which the energy gap E_g between the conduction band and valence band is smaller than the energy gap E_{vc} between (by definition, the top of) the valence band and the outermost core states. While the outermost core states and the conduction band originate from the highest occupied np states and the lowest unoccupied $(n + 1)s$ (and possibly nd also, when $n \geq 3$) states of the cation, respectively, the valence band originates from the highest occupied $n'p$ states of the anion. Due to the Madelung potential with opposite signs for cation and anion sites, the sign of $E_{vc} - E_g$ depends on the combination of cation and anion elements. If $E_{vc} < E_g$, luminescence ($v \rightarrow c$) is the only possible process for the annihilation of the core hole (type L); if $E_{vc} > E_g$, the Auger process (across the gap) competes with and (in most cases) dominates (type A) under the ionizing excitation (core to conduction band). Among the type-L materials are²³ CsF, CsCl, CsBr, RbF and²⁴ BaF₂. The luminescence in these materials is particularly named *Auger-free luminescence* since it is the exceptional case among the recombination luminescence of the core hole in general (the hole in the K shell in heavy elements is also an exception, see Fig. 12.4). Historically, it has also been called *cross luminescence* since it is an interatomic (anion \rightarrow cation) electronic transition.^{25,26}

The Auger-free luminescence in Rb halides has been applied to an ultraviolet-light amplifier with the following merits:²⁷ (1) *any* intensity of ionizing radiation can give rise to the population inversion (zero threshold) since the population inversion between the core states and the valence band takes place as soon as a core hole is produced; (2) since host atoms, instead of impurities, participate in the amplification, the amplification coefficient is much higher than typical solid state lasers, being promising as a nano-sized laser.

12.4 Lattice relaxation of core hole as reflected in recombination luminescence spectra

Deep core holes are in general not long-lived enough to give rise to lattice distortion (which needs 10^{-13} s) because of the recombination luminescence and the Auger processes (see Fig. 12.4 for the case of the K shell), and hence we do not have to consider the effect of lattice relaxation in the emission spectra due to these holes. The annihilation of shallower core holes is in general dominated by Auger processes which, in effects, suppress the observation of luminescence therefrom. However, the Auger-free luminescence mentioned in the preceding section has a sufficient

lifetime ($\sim 10^{-9}$ s due to the small transition energy E , see (3.6.7)) for lattice relaxation to be completed, resulting in the spectral lineshape characteristic of the *core hole* which is quite different^{28,29} from that of the recombination luminescence due to a positive hole or an exciton associated with the *valence band*.

The positive hole created in the core band of an ionic crystal by ionizing radiation will give rise to displacement polarization of the surrounding lattice, thus forming a polaron. Because of the narrow core band and hence the large effective mass, the polaron is supposed to be in the strong coupling regime in which the hole localization and the polarization around it are to be determined self-consistently within the adiabatic picture (see eqs. (9.2.18, 9.2.19)). The polarization gives rise to an electrostatic potential acting back upon the hole which behaves at long distance ($r \gg a_p$, where a_p is the polaron radius) as

$$U(r) \sim -|e|^2 / 4\pi\bar{\epsilon}r, \quad \bar{\epsilon}^{-1} \equiv \epsilon_e^{-1} - \epsilon_s^{-1}, \quad (12.4.1)$$

(see for instance, eq. (9.1.7) derived for the weak coupling case, whose asymptotic behavior at long distance, however, is common to any coupling regime).

During the optical transition of the positive hole to the valence band, this polarization field remains unchanged due to the Franck–Condon principle. Since the effective mass of the hole in the valence band is smaller than in the core band, it is now more shallowly bound by the potential $U(r)$ than it was in the core band.²⁸ Namely, the binding energy b_c of the polaron (negative of the self-energy) within the core band is reduced to $b_v (< b_c)$ after its optical transition to the valence band. As a result, one expects the optical spectrum to consist of a discrete line at $E = E_{cv} - b_v + b_c$ due to the bound-to-bound transition and a continuum at $(E_{cv} - \Delta E_v - b_v < E < E_{cv} - b_v)$ due to the bound-to-continuum transition (where ΔE_v is the width of the valence band). The higher Rydberg states (with lower transition energy) of the former are expected to coalesce and merge into the continuum. This is to be contrasted with the expected spectrum consisting of only a continuum ($E_{cv} - \Delta E_v < E < E_{cv}$) under the neglect of the polaron effect.

Assuming b_v to be ~ 0.5 eV, the two pictures mentioned above are consistent with the known values of E_{cv} and ΔE_v (from band calculations and photoelectron spectroscopy) so far as the expected energy region where the spectra are observed is concerned. The fact that there is always a small peak or hump on the high-energy side of the main peak in Auger-free luminescence spectra lends support to the existence of the polaron effect, although there remain some problems to be confirmed quantitatively, in particular, whether the observed intensity ratio and the binding energy are consistent with theoretical estimations.

A recent time-resolved study of Auger-free luminescence in the femtosecond regime³⁰ clarified that the continuum part of the spectra appears delayed by ~ 200 fs

compared to the discrete part, and this was interpreted in terms of the configuration coordinate model²⁹ based on the above-mentioned polaron picture.

Except in the K-shell excitations of heavy elements ($Z > 30$) and the shallow-core excitations as described above, Auger processes make dominant contributions to the decay rate of the hole. By the Auger process, the hole may be removed to a shallower core while the excited atom becomes charged. Successive (if any) Auger processes finally remove a valence electron which is responsible for the interatomic bonding. It has been supposed that bond breaking and dissociation or desorption of an atom take place after such multistep Auger processes with a multiply-charged state which is naturally unstable.

However, recent studies by SR spectroscopy combined with simultaneous energy analyses of Auger electrons and dissociated ions revealed that the initial motion of atoms immediately after the core excitation can play important roles in the reactions which follow and is even reflected in the Auger electrons. According to the photoelectron spectroscopic study of molecular HBr due to Morin and Nenner,³¹ the resonant excitation of Br-3d core electrons into the lowest unoccupied $4p\sigma$ level resulted in an Auger photoelectron which is unambiguously identified (energetically) with those originating from an isolated $\text{Br}^*(\rightarrow \text{Br}^+ + e^-)$. This Br^* can be produced only after the chemical reaction: $h\nu + \text{HBr} \rightarrow \text{H} + \text{Br}^*$ (superfix * indicates the excited state of the relevant atom).

A more elaborate study of this problem was done on a BF_3 molecule which is of a planar triangular shape with B centrally located on the plane in the ground state, but becomes destabilized, producing lateral movement of B out of the plane in the $1s \rightarrow 2a_2''$ excited state of B. The simultaneous energy analyses of dissociated B^+ and F^- ions and Auger electrons for various excitations strongly indicate that the lateral motion of B commenced during the Auger processes and before the dissociation,³²⁻³⁴ as is in fact reflected in the spectral shape of the Auger electrons.^{35,36}

12.5 Photoelectron spectroscopy

In the optical spectra so far described one can observe only the energy *difference* $\varepsilon_v - \varepsilon_\mu (= \hbar\omega)$ between the initial and final states of matter, whereas in photoelectron spectroscopy³⁷ one can find the energy ε_μ of the initial state *independently* by *direct* measurement of the kinetic energy $\varepsilon (= \hbar^2 k^2 / 2m)$ of the electron emitted (the final state) out of the matter after its photoexcitation through the equation

$$-\varepsilon_\mu = \hbar\omega - \varepsilon (\equiv E_B). \quad (12.5.1)$$

Namely, the depth of the initial state from which electron has been taken can be directly measured as the difference (E_B) of the incident photon energy and the kinetic

energy of the emitted electron. E_B measures the *binding energy* of the electron in the initial state, whence the suffix B. Of course, this simple principle is based on the assumption that the photoexcited electron is subject to no energy loss before it emerges from the matter, which needs preparation of such an idealized situation in the case of solids. Any deviation of ε_μ from its atomic value, such as the chemical shift which provides us with the information on the environment of the atom in the case of inner-shell electrons and on the chemical bonding with surrounding atoms in the case of valence electrons, can be measured by E_B -spectroscopy, as was pioneered by Siegbahn and his group.³⁷ (See Ref. [38] for a general survey of more recent developments.) Using *angle-resolved* photoelectron spectroscopy, one can even find the dispersion $\varepsilon_\mu(\mathbf{k})$ of the initial band by measuring the direction of the photoemission, though within certain restrictions and with further experimental requirements.³⁹

Within the Hartree–Fock approximation, ε_μ denotes the one-electron energy of the occupied state while (12.5.1) represents Koopman’s theorem. Due to the many body effects, however, the hole energy $-\varepsilon_\mu$ is not only shifted by the renormalization effect but is accompanied by simultaneous excitations of modes other than the emitted electron.^{40,41} The most important renormalization effects in solids are the electronic polarization of, and the charge screening by, the surrounding medium in the cases of insulators and metals, respectively. These responses are described in Sections 9.3 and 6.5 in terms of the dielectric constant $\epsilon(\mathbf{k}, \omega)$. Following Kotani,⁴¹ we proceed as follows.

The Hamiltonian of the medium is denoted by H_0 , and its ground state and energy by $|0\rangle$ and E_0 , respectively:

$$H_0|0\rangle = E_0|0\rangle. \quad (12.5.2)$$

Note that we confine ourselves to absolute zero of temperature. By photoexcitation, a core hole with charge $+e$ is suddenly introduced at $\mathbf{r} = 0$, which gives rise to the electrostatic potential $\phi(\mathbf{r}) = e/4\pi\epsilon_0 r$, acting upon the medium with charge density fluctuation

$$\rho(\mathbf{r}) = V^{-1/2} \sum_{\mathbf{q}} \rho_{\mathbf{q}} \exp(i\mathbf{q} \cdot \mathbf{r}) \quad (12.5.3)$$

through the interaction Hamiltonian

$$H_I = \int d\mathbf{r} \rho(\mathbf{r}) \phi(\mathbf{r}) = \sum_{\mathbf{q}} \rho_{\mathbf{q}} \phi_{-\mathbf{q}}, \quad (12.5.4)$$

$$\text{where } \phi_{-\mathbf{q}} \equiv V^{-1/2} \int d\mathbf{r} \phi(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) = V^{-1/2} e/\epsilon_0 q^2. \quad (12.5.5)$$

The eigenstates and eigenenergies of the medium after the photoexcitation are denoted by E_f and $|f\rangle$, respectively, so that

$$H|f\rangle \equiv (H_0 + H_1)|f\rangle = E_f|f\rangle. \quad (12.5.6)$$

In the photoelectron spectroscopy of binding energy, the ordinate, the number of photoelectrons, is in general a function of two variables, the energy $\hbar\omega$ of the incident photon *and* the kinetic energy ε of the photoelectron, and not a function of a single parameter $E_B = \hbar\omega - \varepsilon$. Apart from an unimportant factor which depends gently on each of them, however, the photoelectron spectrum is considered to be given by

$$F(E_B) = \sum_f \langle 0|f\rangle|^2 \delta(\varepsilon - \varepsilon_\mu + E_f - E_0 - \hbar\omega), \quad (12.5.7)$$

which is rewritten in the Fourier transform of the generating function $f(t)$ introduced in (4.5.10) as

$$\begin{aligned} F(E_B) &= \langle 0|\delta(\varepsilon - \varepsilon_\mu + H - E_0 - \hbar\omega)|0\rangle \\ &= (1/\pi\hbar)\text{Re} \int_0^\infty dt f(t) \exp[i(E_B + \varepsilon_\mu)t/\hbar], \end{aligned} \quad (12.5.8)$$

where

$$f(t) = \langle 0|\exp[-i/\hbar](H - E_0)t|0\rangle = \langle 0|L(t)|0\rangle, \quad (12.5.9)$$

$$L(t) \equiv \exp[(i/\hbar)H_0t] \exp[-(i/\hbar)(H_0 + H_1)t]. \quad (12.5.10)$$

Under the assumption of a vanishing expectation value of the charge density in the ground state of matter, namely

$$\langle 0|\rho_q|0\rangle = 0 \quad \text{for all } \mathbf{q}, \quad (12.5.11)$$

we have $\langle 0|H_1|0\rangle = 0$, and hence obtain the cumulant expansion

$$\langle 0|L(t)|0\rangle = \exp \left[-\hbar^{-2} \int_0^t d\tau (t - \tau) \langle 0|H_1(\tau)H_1|0\rangle + (\text{higher-order cumulants}) \right], \quad (12.5.12)$$

by making use of the derivation of (11.4.11a) from the definition (11.4.5) where $S(t)$ is rewritten as $L(t)$ in the present case.

The calculation of the correlation function in the second-order cumulant reduces, because of (12.5.4), to that of the charge density which is obtained from the inverse transform of (6.1.34):

$$\begin{aligned} \langle 0|\rho_q(\tau)\rho_{-q}|0\rangle &= -(\hbar q^2/\pi) \times \int_{-\infty}^\infty d\omega \exp(-i\omega\tau) \\ &\times [1 - \exp(-\beta\hbar\omega)]^{-1} \text{Im}[\chi_{\parallel}(\mathbf{q}, \omega)]. \end{aligned} \quad (12.5.13)$$

This is an example of the fluctuation–dissipation theorem, applied to the charge density fluctuation versus dielectric loss. At absolute zero of temperature ($\beta \rightarrow +\infty$) to which we confine ourselves, only the region of positive ω contributes to the integral. Inserting (12.5.4) and (12.5.13) into (12.5.12) and integrating over time, we obtain, neglecting the higher-order cumulants,

$$\langle 0|L(t)|0\rangle = \exp[-S - (it/\hbar)\Delta + S(t)] \quad (12.5.14)$$

where

$$S(t) \equiv \int d\omega s(\omega) \exp(-i\omega t), \quad S \equiv S(0) = \int d\omega s(\omega), \quad (12.5.15)$$

$$\Delta \equiv - \int d\omega s(\omega) \hbar \omega, \quad (12.5.16)$$

$$s(\omega) \equiv (\pi\hbar)^{-1} (e/\epsilon_0)^2 V^{-1} \sum_{\mathbf{q}} q^{-2} \omega^{-2} \text{Im}[\chi_{\parallel}(\mathbf{q}, \omega)]. \quad (12.5.17)$$

By expanding (12.5.14) in powers of $S(t)$, one finds that the photoelectron spectrum (12.5.8) starts with the zero-line at $E_B = -\epsilon_\mu + \Delta$, followed by its successive sidebands (satellites) given by the convolutions of the fundamental spectrum $s(\omega)$ as shown schematically in Fig. 12.5, the same as the zero-phonon line and its phonon

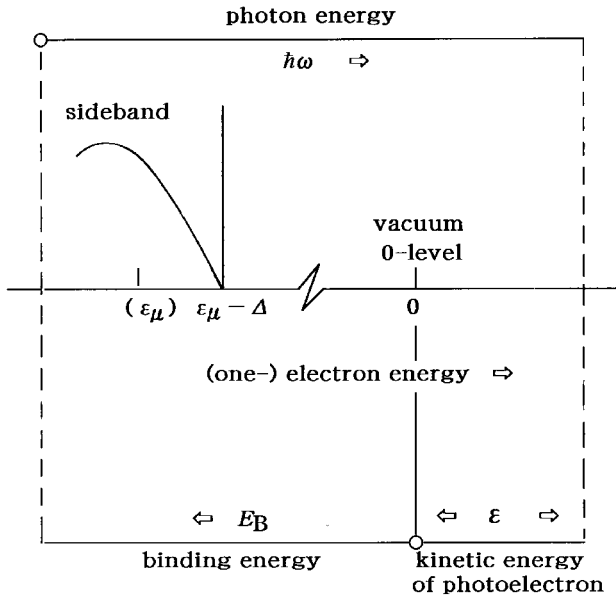


Fig. 12.5 Schematic representation of photoelectron spectra of a core state μ with self-energy shift, accompanied by sidebands due to simultaneous excitations. The abscissa of the spectra is the binding energy E_B (increasing toward the left along the bottom line), which is the incident photon energy $\hbar\omega$ minus the kinetic energy ϵ of the emitted electron.

sidebands of the optical absorption spectrum of a localized electron described in Section 4.5:

$$F(E_B) = \sum_{n=0}^{\infty} F_n(E_B), \quad (12.5.18)$$

$$F_0(E_B) = \exp(-S)\delta(E_B + \varepsilon_\mu - \Delta), \quad (12.5.19)$$

$$\begin{aligned} F_n(E_B) = \exp(-S)(n!)^{-1}\hbar^{-1} \int d\omega_1 \int d\omega_2 \dots \int d\omega_{n-1} \\ \times s([\{E_B + \varepsilon_\mu - \Delta\}/\hbar] - \omega_1 - \omega_2 \dots - \omega_{n-1}) \\ \times s(\omega_1)s(\omega_2) \dots s(\omega_{n-1}). \end{aligned} \quad (12.5.20)$$

The integrated intensities obey the Poisson distribution, $I_n = \exp(-S)S^n/n!$, and $I_0 = \exp(-S)$ represents the square of the overlap integral between the ground states of H_0 and $H_0 + H_I$.

One can rewrite the energy Δ , defined by (12.5.16) and (12.5.17), with the use of the dispersion relation (6.2.2) with $\omega = 0$ applied to the odd function $\text{Im } \chi(\mathbf{q}, \omega)$ and of the relation (6.3.3) with ϵ_e being replaced by ϵ_0 (note that we are now concerned with the contribution of *electronic* polarization to the dispersion of $\chi(\mathbf{q}, \omega)$ in the background of a *vacuum*), as

$$\begin{aligned} \Delta &= -V^{-1} \sum_{\mathbf{q}} (e^2/2q^2) [\epsilon_0^{-1} - \epsilon_e(\mathbf{q}, 0)^{-1}] \\ &= -(1/2) \sum_{\mathbf{q}} [\epsilon_0^{-1} - \epsilon_e(\mathbf{q}, 0)^{-1}] \mathbf{D}_{\mathbf{q}}^2 \end{aligned} \quad (12.5.21)$$

where $\mathbf{D}_{\mathbf{q}} = V^{-1/2}e\mathbf{q}/q^2 = V^{-1/2}\epsilon_0\mathbf{q}\phi_{\mathbf{q}}$ is the Fourier component of the electric flux density $\mathbf{D}(\mathbf{r}) = -\epsilon_0\nabla\phi(\mathbf{r}) = e\mathbf{r}/r^3$ due to the positive charge e at $\mathbf{r} = 0$. Here Δ means the energy *decrease* (*stabilization* energy, or *relaxation* energy) due to the introduction of electric charge e from a vacuum with dielectric constant ϵ_0 into a medium with *electronic* dielectric constant $\epsilon_e(\mathbf{q}, \omega)$. According to (12.5.8, 12.5.14), Δ appears in the photoelectron spectrum in the form of $(\epsilon\mu - \Delta)$, namely, the energy of the core hole (the depth), $-\varepsilon_\mu (> 0)$, is diminished (stabilized) by $|\Delta|$ due to the electronic polarization of the medium, as shown schematically in Fig. 12.5.

In insulators, one can put $\mathbf{q} = 0$ in the susceptibility χ since its spatial dispersion is not so important, and obtain

$$\Delta = -(e^2k_0/4\pi)(k_0/\pi)(\epsilon_0^{-1} - \epsilon_e^{-1}) \quad (12.5.22)$$

which is in agreement with (9.3.23), the classical limit (appropriate since we are neglecting the dispersion of the core band) of the electronic-polaron energy with sign reversed (since we are concerned with the energy of a hole – *anti*-particle, instead of an electron – particle). Here k_0 is the Debye cut-off wave vector q

defined by

$$N = \sum_q (V/8\pi^3) 4\pi \int_0^{k_0} q^2 dq, \quad \text{namely} \quad k_0 = (6\pi^2/v_0)^{1/3}. \quad (12.5.23)$$

Equation (12.5.18) represents the self-energy of a charged particle with radius $a = \pi/2q_0$ embedded in a dielectric continuum with dielectric constant $\bar{\epsilon}$ given by $\bar{\epsilon}^{-1} = \epsilon_0^{-1} - \epsilon_c^{-1}$, being consistent with (12.5.19).

In metals, one puts dielectric constant (6.5.8) into (12.5.18) and obtains

$$\Delta = -(1/2)(e^2/4\pi\epsilon_0)(\pi/2)q_s \tan^{-1}(k_0/q_s) \quad (12.5.24)$$

where $\hbar q_s$ is the Fermi momentum. Both (12.5.22) and (12.5.24) are of the order of eV.

In insulators, the first-order satellite, given by $s(\{E_B + \epsilon_\mu - \Delta\}/\hbar)$, consists of a series of simultaneously emitted longitudinal excitons (the *valence* exciton consisting of a positive hole in the valence band and an electron in the conduction band). However, the valence exciton produced around the core hole is expected to form a state deeply bound to the core hole. In fact, the electron is subject to the doubled attractive potential due to the valence-band hole and the core hole, resulting in four times as large (in the optimum estimation) a binding energy of the valence exciton and hence in a binding energy of core hole–valence exciton complex (CHVXC) nearly three times as large, which is quite significant. Probably, the fractional intensity $\exp(-S)S$ of the satellite is mostly concentrated to this complex. However, there seems to be no report so far on such a structure. A similar mechanism has been studied with the bound states of an optical phonon with a valence exciton (EPBS).^{20,21} In order to estimate the coupling strength S governing the intensities of the satellites, let us make a simplifying assumption that the spectrum of $s(\omega)$ is concentrated at the energy ω_{ex} of a typical valence exciton or its complex with the core hole. In view of (12.5.15, 12.5.16), one can write

$$s(\omega) = S\delta(\omega - \omega_{\text{ex}}), \quad S = \Delta/\hbar\omega_{\text{ex}}. \quad (12.5.25)$$

Since both Δ and $\hbar\omega_{\text{ex}}$ are of the order of several electron-volts in wide-gap insulators, the coupling strength S is of the order of unity and the higher-order satellites with intensities given by the Poisson distributions will not be so significant.

In metals, we have plasma resonance due to the *collective* motion of electrons

$$\text{Im}\chi(0, \omega) = (\pi/2)\epsilon_0\omega_p\delta(\omega - \omega_p)$$

according to (6.5.6, 6.5.7), and

$$s(\omega) = (e^2/4\pi\epsilon_0)(k_0/\pi)(\hbar\omega_p)^{-1}\delta(\omega - \omega_p). \quad (12.5.26)$$

Namely, the plasma satellites are expected to appear at $E_B = \epsilon_\mu + \Delta + n\hbar\omega_p$ with the fractional intensities $I_n = \exp(-S)S^n/n!$. Since $\hbar\omega_p$ is of the order of several electron-volts, at most one plasmon satellite is expected, as has in fact been observed in 2p core photoelectron spectra of metallic Na.⁴² On the other hand, the low-energy excitations of *individual* electrons make divergent contribution $s(\omega) \propto \omega^{-1}$ ($\omega \rightarrow 0$) with $S \rightarrow \infty$ due to the finite state density at the Fermi energy. It is not so simple to describe this problem with the use of the dielectric function as we were able to do for the calculation of the energy shift Δ , and we will resort to a simplified Hamiltonian in the next section.

Finally, we have to take into account the broadening effect of the core level by the convolution, into the photoelectron spectra obtained above, of the broadening functions such as the Gaussian band due to the fluctuating phonon field (Section 4.3) and the Lorentzian band of lifetime broadening due to radiative annihilation and the Auger process (Section 12.3). While the interactions with phonons and photons are not included in the interaction H_I (eq. (12.5.4)) with other electrons, one might suspect undue double counting of the interelectron interaction, one in the photoelectron spectra described above and one in the Auger lifetime. However, the Auger process needs other electrons to annihilate the core hole concerned, while no such process is included in H_I because other electrons are treated simply as a dielectric medium, being distinguished from the core electron. This situation has some resemblance to the zero-phonon line (Section 4.5) which was described with the same formalism as the present one, and its broadening is caused by (i) radiative de-excitation and (ii) nonradiative multiphonon de-excitation through the interstate matrix elements of the electron-phonon interaction not considered in Chapter 4.

12.6 Many-body effects in normal metals

We have described the behaviors of the core hole, as reflected in the photoelectron spectra, in terms of the dielectric theory in which the long-range part of the Coulomb interactions among the electrons in the medium is taken into account through the effective electric field as described in Section 6.3 (to be more specific, through the relations (6.3.2) to (6.3.4) with ϵ_e replaced by ϵ_0). In metals, collective motions such as plasma oscillations are determined by the long-range Coulomb interactions, while the individual electrons seem to behave as if they are almost unaware of their mutual interactions, as has been recognized throughout the historical success of the free-electron model (Drude-Sommerfeld theory) in describing the thermal and electrical properties of metals.⁴³ It is now well understood⁴⁴ that

as a result of introducing collective excitation modes the interactions between individual electrons reduce effectively to the screened, short-ranged interactions as given by (6.5.11). This justifies the use of the effective Hamiltonian H_{eff} with weak or no interactions as mentioned in Section 6.3, at least in describing the *individual* excitations of electrons. For simplicity, we consider electrons with *no mutual* interactions but subject to short-range *on-site* interaction $-v$ with the core hole. Namely, we put

$$H_0 = \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} a_{\mathbf{k}}^+ a_{\mathbf{k}}, \quad H_1 = -(v/N) \sum \sum_{\mathbf{k} \neq \mathbf{k}'} a_{\mathbf{k}}^+ a_{\mathbf{k}}'. \quad (12.6.1)$$

The diagonal terms with $\mathbf{k} = \mathbf{k}'$, which have been omitted from H_1 in (12.6.1) to justify the use of cumulant expansion (12.5.12), contribute the first-order perturbation energy $\Delta_1 = -v$ in the case of a monovalent metal ($N_e = N$). This on-site interaction is to be defined, with the use of (6.5.10, 6.5.11), by

$$-v \equiv -v_0^{-1} \int d\mathbf{r} (e^2/4\pi\epsilon_0 r) \exp(-q_s r) = -(2\varepsilon_F/3) \quad (12.6.2)$$

where v_0 is the volume of a unit cell, q_s the reciprocal of the screening length and ε_F the Fermi energy. In normal metals, this first-order energy Δ_1 is already below Δ of (12.5.25) which is a legitimate value. The discrepancy gets greater if the second-order perturbation energy Δ (<0) to be obtained through the following description is added to Δ_1 , but would be diminished by considering the increase in the so far neglected electron-electron interaction energy due to their charge redistribution by the core hole. It is difficult to estimate and explain the still remaining discrepancy (if any) since we are resorting here to the different Hamiltonian (12.6.1). Nevertheless, we assume it to be enough for qualitative description of the divergent behaviors of the spectra in metals with which we shall be concerned in this section.

Making use of the Heisenberg representation

$$\begin{aligned} H_1(\tau) &= \exp(i\tau H_0/\hbar) H_1 \exp(-i\tau H_0/\hbar) \\ &= -(v/N) \sum \sum_{\mathbf{k} \neq \mathbf{k}'} \exp(i\tau \varepsilon_{\mathbf{k}}/\hbar) a_{\mathbf{k}}^+ a_{\mathbf{k}'} \exp(-i\tau \varepsilon_{\mathbf{k}'}/\hbar), \end{aligned}$$

one can write

$$\langle 0 | H_1(\tau) H_1 | 0 \rangle = (v/N)^2 \sum_{\mathbf{k} > k_F} \sum_{\mathbf{k}' < k_F} \exp[-i(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}'})\tau/\hbar]$$

where $|0\rangle$ is the ground state of H_0 in which electrons occupy the Fermi sphere with radius k_F . Thus one obtains the generating function (12.5.15) with (12.5.16, 12.5.17) where

$$\begin{aligned} s(\omega) &\equiv (v/N)^2 \omega^{-2} \sum_{\mathbf{k} > k_F} \sum_{\mathbf{k}' < k_F} \delta[\omega - (\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}'})/\hbar] \\ &= \alpha \omega^{-1} (\hbar \omega \ll \varepsilon_F), \quad \alpha \equiv (\rho v)^2. \end{aligned} \quad (12.6.3)$$

Here ρ is the density of state per unit volume of electrons which is assumed to be constant within $\hbar\omega$ around the Fermi energy ε_F .

The factor ω^{-1} in (12.6.3), which originates from the finite density of state at the Fermi energy and hence is characteristic of all metals, gives rise to the infrared ($\omega \rightarrow 0$) divergence of the integral S defined by (12.5.16). According to the statement just below (12.5.21), this means that the ground states of $H_0 + H_I$ and H_0 are orthogonal to each other, indicating the serious effect of charge redistribution caused by a positive charge introduced in metals, as it is called “orthogonal-ity catastrophe”.⁴⁵ The infrared divergence, $s(\omega) \sim \omega^{-1}$, causes the singular line-shapes of photoelectron spectra as well as photoabsorption spectra in metals, as has been extensively studied by Anderson,⁴⁵ Mahan,⁴⁶ Nozières and De Dominicis,⁴⁷ Hopfield,⁴⁸ Friedel,⁴⁹ Doniach and Sunjic,⁵⁰ and many others.

We now describe the singularity of the photoelectron spectra with the use of the infrared divergent $s(\omega)$ given by (12.6.3). Denoting the upper cut-off energy $\hbar\omega$ in (12.6.3) by D , one obtains the generating function as

$$f(t) = \exp[-it\Delta/\hbar - \alpha \ln(itD/\hbar)], \quad (12.6.4)$$

and the photoelectron spectrum as

$$F(E_B) = \begin{cases} 0, & (E_B + \varepsilon_\mu - \Delta < 0) \\ [D\Gamma(\alpha)]^{-1}[(E_B + \varepsilon_\mu - \Delta)/D]^{-(1-\alpha)}, & (E_B + \varepsilon_\mu - \Delta > 0) \end{cases} \quad (12.6.5)$$

where $\Gamma(\alpha)$ is the Gamma function. The expectation value of the number of electron-hole pairs excited simultaneously with the core-electron excitation is given by $\langle \bar{n} \rangle = S = \infty$ due to the divergent contribution of low-energy excitations.

If one makes use of a more exact treatment of the singularity than the first-order cumulant expansion (see (12.5.13)) and takes account of the spin degeneracy, the α in eqs. (12.6.4, 12.6.5) is to be replaced by

$$\alpha = 2 \sum_{\ell} (2\ell + 1) [\delta_{\ell}(\varepsilon_F)/\pi]^2, \quad (12.6.6)$$

where $\delta_{\ell}(\varepsilon_F)$ is the phase shift of the partial wave with angular momentum ℓ of the conduction electron at the Fermi energy due to the core-hole potential v .

In connection with this, we should mention the Friedel sum rule concerning the phase shift $\delta_{\ell}(\varepsilon_F)$.⁵¹ Suppose a core hole is suddenly created at time $t = 0$. Then the conduction electrons are scattered by its potential and thereby redistribute so as to screen its charge. If one disregards the finite lifetime of the core hole, the system tends to its lowest relaxed state $|g^h\rangle$ in the limit of $t \rightarrow \infty$. The effect of screening is represented by the phase shift $\delta_{\ell}(\varepsilon)$ of the conduction electrons, which appear in the asymptotic form of the radial wave function

$$u_{k\ell}(r) = \sin[kr + \delta_{\ell}(\varepsilon) - \ell\pi/2]. \quad (12.6.7)$$

With the use of $u_{k\ell}(r)$ one can calculate the change of the conduction electron number around the core hole, which should be unity for perfect screening of the hole charge:

$$(2/\pi) \sum_{\ell} (2\ell + 1) \delta_{\ell}(\varepsilon_F) = 1. \quad (12.6.8)$$

This is the Friedel sum rule. It gives a certain constraint on the value of α given by (12.6.6).

According to Anderson,⁴⁵ the lowest states of the Fermi sea, $|g^{\downarrow}\rangle$ and $|g\rangle$ in the presence and absence of the core hole, respectively, are orthogonal; more precisely, the overlap integral $\langle g_N | g^{\downarrow} N \rangle$ in the N -conduction-electron system is expressed as

$$\ln |\langle g_N | g_N^{\downarrow} \rangle| = - \sum_{\ell} (2\ell + 1) [\delta_{\ell}(\varepsilon_F)/\pi]^2 \ln N, \quad (12.6.9)$$

so that

$$\lim_{N \rightarrow \infty} \langle g_N | g_N^{\downarrow} \rangle = \langle g | g^{\downarrow} \rangle = 0.$$

A singularity similar to that of the photoelectron spectrum appears also in the photoabsorption spectrum corresponding to the excitation of core electrons to the vacant states above the Fermi sea of the conduction band. We give here only the result. With the absorption edge given by $E_t = \varepsilon_F - \varepsilon_c - (E_{g^{\downarrow}} - E_g)$, the absorption spectrum of a core hole with orbital angular momentum ℓ is given by

$$F_a(E) = F_{\ell-1}(E) [D/(E - E_t)]^{\alpha_{\ell-1}} + F_{\ell+1}(E) [D/(E - E_t)]^{\alpha_{\ell+1}}, \quad (12.6.10)$$

where $F_{\ell \pm 1}(E)$ is the spectrum for vanishing v , and the ℓ -dependent exponent is given by

$$\alpha_{\ell} = (2\delta_{\ell}/\pi) - 2 \sum_{\ell'=0}^{\infty} (2\ell' + 1) (\delta_{\ell'}/\pi). \quad (12.6.11)$$

This is the result derived originally by Nozières and De Dominicis.⁴⁷ The spectrum (12.6.10) is asymptotically exact when E approaches E_t . If $\alpha_{\ell-1}$ or $\alpha_{\ell+1}$ is positive, $F_a(E)$ is enhanced and diverges at the Fermi edge, but if both of them are negative, $F_a(E)$ is suppressed and becomes a rounded edge. In the same way, the luminescence spectrum also has a Fermi edge singularity with the same exponent as that of the absorption spectrum. Some extensions of the result (12.6.10) have also been made by taking account of the electron-hole exchange^{52,53} and the coexistence of exchange and spin-orbit interaction,⁵⁴ and by extending the validity range of E with the use of spectral convolution.⁵⁵

According to experimental studies,⁵⁶⁻⁵⁸ the spectra at the L_{23} edges of Na, Mg and Al are sharp and peaked, while those of the K edges of Li and Al are broad and rounded. If one assumes that the screened potential of the core hole is sufficiently

short ranged, s-wave scattering is predominant, i.e. $\delta_0 \gg \delta_1, \delta_2$, etc., so that one obtains $\alpha_0 > 0$ and $\alpha_1 < 0$ from (12.6.11), together with the Friedel sum rule (12.6.8). This means that the L_{23} edge should be peaked and the K edge rounded, in qualitative agreement with the experimental data. However, it has been a subject of controversy whether those line shapes are qualitatively reproduced by the Mahan–Nozières–De Dominicis theory alone. It was pointed out by Dow (see Ref. [59] and references therein) and others⁶⁰ that the electron–phonon interaction is important in the K edge spectrum of Li, and that one-electron characteristics, both effects of band density of states and transition matrix element, are essential for the L edge of Mg and K edge of Al.

12.7 Core-electron spectra of atoms with incomplete outer shells

As we look at the electronic configurations of all atoms which constitute the rows and columns of the periodic table, we first note the general rule that the electrons occupy one-electron energy levels from the lowest subshell to higher ones: 1s, 2s, 2p, 3s, 3p, 4s, 3d, 4p, 5s, 4d, 5p, 6s, 5d, 4f, 6p, 6d, 5f. A remarkable fact, found in the groups of the so-called *transition elements*, is that with the increase of atomic number Z the number of electrons occupying the nd or nf subshell with well-localized orbitals increases one by one without affecting the occupation numbers of the spatially outer shells (apart from a few irregularities within the sets): (4s, 3d), (5s, 4d), (5d, 4f), (6s, 5d), and (6d, 5f) (the latter in each set is more localized while the former has a lower one-electron energy). This means that deepening of the Coulomb potential due to an increase in Z is about enough to accommodate another electron in the nd or nf subshell provided it is not yet completely filled up (*incomplete* subshell), keeping the occupation of the outer subshell(s) unchanged because the localized orbital is subject to much larger energy lowering than the extended ones. It indicates that an increase of the nuclear charge by one is nearly compensated by the increase of one electron in the inner subshell, leaving no significant effect on the outer shell, and hence no change in the chemical properties of the elements. This is the well-known explanation why several elements (transition elements) with successive atomic number Z belong to one compartment (column-row) in the periodic table.

In the condensed state, these transition elements are mostly metallic with a conduction band formed from the outer subshells mentioned above. The inner subshell remains incomplete as it was in the atom, with no further electrons flowing in from the conduction band. This fact indicates a large Coulomb repulsion between localized electrons which prevents such an inflow. A simple concept of the Fermi level based on the one-electron energy scheme applies to the conduction band but not to the inner subshell which is subject to strong electron correlation.

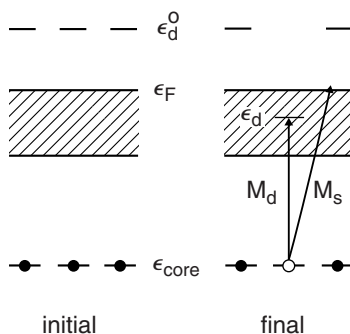


Fig. 12.6 Initial and final states in the optical excitation of a core electron, with unoccupied d level ϵ_d pulled down below the Fermi level ϵ_F .⁶¹

This argument leads us to imagine that the core-electron excitation of metals with an incomplete subshell (nd or nf , which will hereafter be referred to as d shell, representatively), being in effect equivalent to an increase in the nuclear charge, causes one of the conduction electrons to be pulled into the incomplete subshell of the excited atom. It is natural to imagine that the energy level of the d orbital to be filled up *has been* at ϵ_d^0 well above the Fermi level ϵ_F of the conduction electrons *before* the core-electron excitation, but *is now pulled down* to ϵ_d due to the core hole, *possibly below* ϵ_F , as shown schematically in Fig. 12.6. However, the d orbital concerned is different from the orbital constituting the conduction band (denoted by s orbital for convenience), and we need to consider the problem of the relative probabilities of *two independent events*: a conduction electron *plunges* into the pulled-down d level *or not*.⁶¹

If not, a hole is left behind at ϵ_d within the Fermi sea which is naturally subject to subsequent annihilation by a conduction electron at a finite rate $1/\tau$. One might imagine that as in normal metals the core electron can be excited only to the vacant level above ϵ_F in the conduction band, forming a sharp absorption edge (step-like but for the infrared divergence) at $E \sim \epsilon_F - \epsilon_c$. However, this sharp edge is in fact blurred by \hbar/τ due to the above-mentioned lifetime of the hole in the Fermi sea as the final state of the many-electron system, as shown schematically in region (A-2) of Fig. 12.7.

The real threshold appears below this blurred edge, apparently contrary to the Pauli exclusion principle within the one-electron picture. We will show this by considering the first alternative case mentioned above in which the pulled-down d level *is* occupied by an electron in the final state of core-electron excitation. From the many-electron point of view, it is meaningless to ask whether the d electron came from the core level or the conduction band. Anyway, the lowest possible energy of the final state consists of one electron in the pulled-down d level at ϵ_d , and the

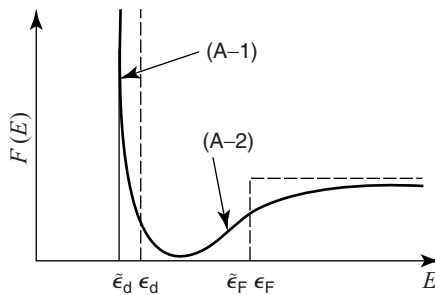


Fig. 12.7 Schematic lineshape of absorption spectra in the case of $\varepsilon_d < \varepsilon_F$. The spectrum in the absence of s-d mixing is shown with dashed lines.⁶¹

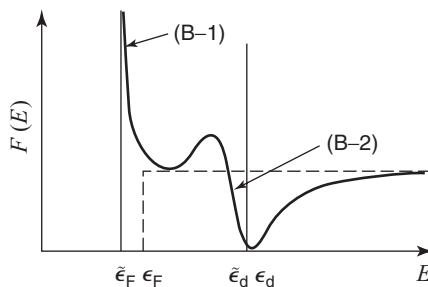


Fig. 12.8 Schematic lineshape of absorption spectra in the case of $\varepsilon_d > \varepsilon_F$. The spectrum in the absence of s-d mixing is shown with dashed lines.⁶¹

other electrons in the conduction band as before the excitation. The net energy difference between this final state and the initial state before excitation is given by $E \sim \varepsilon_d - \varepsilon_c (< \varepsilon_F - \varepsilon_c$ as mentioned above), as shown schematically in region (A-1) of Fig. 12.7. This threshold shows a singularity due to the infrared divergence (as mentioned in Section 12.6) of simultaneous excitations of conduction electrons scattered by the localized d level which has suddenly emerged.

One can thus expect two characteristic portions in the photoabsorption spectra: the *real* threshold with singularity (A-1) of the many-electron system, and the *blurred* Fermi edge (A-2) of the one-electron picture.

The above argument is confined to the case (A) of $\varepsilon_d < \varepsilon_F$. It is instructive to consider the opposite case (B) of $\varepsilon_d > \varepsilon_F$ and compare both results. In this case, the Fermi edge at $E \sim \varepsilon_F - \varepsilon_c$ is the real threshold, with the singularity of infrared divergence again as shown in region (B-1) of Fig. 12.8. The d level at $E \sim \varepsilon_d - \varepsilon_c$ shows up as the Fano resonance with a dip on one side as shown in region (B-2) of the same figure (an example of the Fano effect was presented in Section 10.2, see the asymmetric Lorentzian lineshape of (10.2.12)) which is characteristic of a discrete line embedded in the continuum.

The conceivably simplest model to describe this situation is to ignore the spin and d-orbital degeneracies, and to assume that the d level before and after the core excitation is at $\varepsilon_d^0 (\rightarrow \infty)$ far above the Fermi level ε_F and at ε_d , respectively.⁶¹ Before photoexcitation, we consider simply the system of conduction electrons with Hamiltonian

$$H^{(0)} = \sum_k \varepsilon_k a_k^\dagger a_k, \quad (12.7.1)$$

which is Fermi degenerate in the ground state with energy E_g :

$$|g\rangle = \prod_k^{(\varepsilon_k < \varepsilon_F)} a_k^\dagger |\text{vac}\rangle. \quad (12.7.2)$$

After photoexcitation, we can write the Hamiltonian as

$$H = H_0 + H', \quad (12.7.3)$$

$$H_0 = \sum_k \varepsilon_k a_k^\dagger a_k + \varepsilon_d a_d^\dagger a_d, \quad (12.7.4)$$

$$H' = V \sum_k (a_k^\dagger a_d + a_d^\dagger a_k), \quad (12.7.5)$$

where V is the s–d mixing interaction under the core-hole potential.

Neglection of the d level in (12.7.1, 12.7.2) is justified in the limit: $\varepsilon_d^0 \rightarrow \infty$, while the effect of the core hole is taken into account by the *emergence of the d level* in (12.7.4) rather than simply by an *attractive potential*. Neglection of a core hole in (12.7.3) to (12.7.5) makes the total number of electrons before and after photoexcitation to be N and $N + 1$, respectively. Hence, one can write the optical excitation operator with the use of the creation operators only, as

$$M^\dagger = M_s \sum_k a_k^\dagger + M_d a_d^\dagger, \quad (12.7.6)$$

neglecting the k dependence of M_s , and the absorption spectra as

$$F(E) = \sum_f \langle f | M^\dagger | g \rangle^2 \delta(E - E_f + E_g + \varepsilon_c), \quad (12.7.7)$$

where the summation is to be extended over all eigenstates $|f\rangle$ of H . One can rewrite this as

$$\begin{aligned} F(E) &= \langle g | M \delta(E + \varepsilon_c - H + E_g) M^\dagger | g \rangle \\ &= \pi^{-1} \text{Im} \langle g | M (z - H + E_g)^{-1} M^\dagger | g \rangle \\ &= \pi^{-1} \text{Im} \left[M_s^2 \sum_{k>} \sum_{k'>} G_{kk'} + M_s M_d \sum_{k>} (G_{kd} + G_{dk}) + M_d^2 G_{dd} \right], \end{aligned} \quad (12.7.8)$$

where the summation for the suffix $k >$ is to be extended on $k > k_F$, and

$$z \equiv E + \varepsilon_c - i\eta \quad (\eta \rightarrow +0), \quad (12.7.9)$$

$$G_{kk'} \equiv \langle g | a_k \{z - H + E_g\}^{-1} a_{k'}^\dagger | g \rangle, \text{ etc.} \quad (12.7.10)$$

The summations containing the G s in (12.7.8) are not independent but satisfy the relations:

$$G_{dd} = G_d^0 + (V G_d^0)^2 \sum_{k>} \sum_{k'>} G_{kk'}, \quad (12.7.11)$$

$$\sum_{k>} G_{kd} = \sum_{k>} \sum_{k'>} G_{kk'} V G_d^0 = \sum_{k>} G_{dk}, \quad (12.7.12)$$

where

$$G_d^0 = \langle g | a_d [z - H_0 + E_g]^{-1} a_d^\dagger | g \rangle = [z - \varepsilon_d]^{-1}. \quad (12.7.13)$$

Equations (12.7.11) and (12.7.12) follow immediately from the operator identities:

$$\begin{aligned} [Z - H]^{-1} &= [Z - H_0]^{-1} + [Z - H]^{-1} H' [Z - H_0] \\ &= [Z - H_0]^{-1} + [Z - H_0]^{-1} H' [Z - H_0]^{-1} \\ &\quad + [Z - H_0]^{-1} H' [Z - H]^{-1} H' [Z - H_0]^{-1} \quad (Z \equiv z + E_g) \end{aligned}$$

if one takes appropriate matrix elements on both sides of them.

We will now study the analytic features of absorption spectra for various situations.

Case (A) $\varepsilon_d < \varepsilon_F$

(A-1) Spectral region: $E \sim \bar{\varepsilon}_d - \varepsilon_c$ where $\bar{\varepsilon}_d = \varepsilon_d + \Delta_d$ and Δ_d is the self-energy of the state $a_d^\dagger |g\rangle$ due to H' . With the use of (12.7.11, 12.7.12) one can write $F(E)$ in the form

$$F(E) = [\pi^{-1} \text{Im} G_{dd}] [M_d + M_s(E - \varepsilon_d + \varepsilon_c)/V]^2. \quad (12.7.14)$$

In calculating the first factor of the r.h.s. of this equation it is convenient to calculate its Fourier transform, i.e., the so-called generating function (see eq. (4.5.10)):

$$\begin{aligned} f_d(t) &= \int_{-\infty}^{+\infty} dE \exp(-iEt/\hbar) [\pi^{-1} \text{Im} G_{dd}] \\ &= \langle g | a_d \exp[-i(H - E_g)t/\hbar] a_d^\dagger | g \rangle \\ &= \langle g | a_d \exp[-i(H_0 - E_g)t/\hbar] S(t) a_d^\dagger | g \rangle \\ &= \exp(-i\varepsilon_d t/\hbar) \langle d | S(t) | d \rangle, \end{aligned} \quad (12.7.15)$$

where $|d\rangle \equiv a_d^\dagger |g\rangle$ is the ground state of H_0 , and

$$S(t) \equiv \exp[iH_0 t/\hbar] \exp[-iHt/\hbar] \quad (12.7.16)$$

is of the same form as (12.5.10) and (11.4.5). The latter can be expanded in an infinite series of integrals as in (11.4.9) which can symbolically be written in the

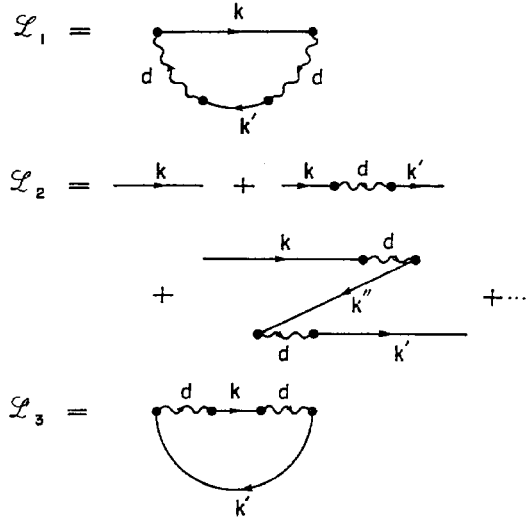


Fig. 12.9 Diagrams making contributions to the most divergent terms. The solid line and the wavy line represent propagators associated with s and d states, respectively. The vertices (solid circles) mean the s-d mixing interaction.⁶¹

time-ordered exponential:

$$S(t) = T \exp \left[-i \int_0^t d\tau H'(\tau)/\hbar \right], \quad H'(\tau) \equiv \exp[i H_0 \tau / \hbar] H' \exp[-i H_0 \tau / \hbar]. \quad (12.7.17)$$

The time-ordering operator T indicates that in the general terms in the infinite series expansion (11.4.9) the integral variables should be ordered such that $\sigma_1 > \sigma_2 > \dots$, as is in fact the case.

According to the linked cluster expansion theorem, one can write

$$\langle d | S(t) | d \rangle = \exp[L_1(t)], \quad (12.7.18)$$

where $L_1(t)$ is the sum of all the connected diagrams in the perturbation expansion of the l.h.s. (similar to (11.4.9)). Assuming that $\rho V^2 \ll (\varepsilon_F - \varepsilon_d)$ where ρ is the density of states of the s band, one calculates $L_1(t)$ in the most divergent term approximation (so as to be asymptotically correct at $|t| \rightarrow \infty$). It is found that the most divergent term appears only in the diagram shown in Fig. 12.9, so that $L_1(t)$ is expressed as

$$\begin{aligned} L_1(t) &\sim -i\Delta_d t + V^4(\varepsilon_d - \varepsilon_F)^{-2} \sum_{k>} \sum_{k'<} [\exp\{-i(\varepsilon_k - \varepsilon_{k'})t\} - 1](\varepsilon_k - \varepsilon_{k'})^{-2} \\ &= -i\Delta_d t - g'^2 \ln(iDt) \quad (|t| \rightarrow \infty), \end{aligned} \quad (12.7.19)$$

$$\text{where } g' \equiv \rho V^2 / (\varepsilon_d - \varepsilon_F), \quad (12.7.20)$$

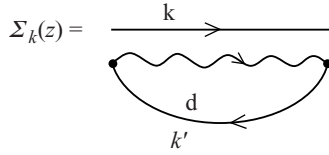


Fig. 12.10 Diagram for the self-energy correction with d hole.⁶¹

and D is the half-width of the s band. The term $-i\Delta_d t$ representing the self-energy shift of the state $|d\rangle$ has been gathered from all orders of the perturbation expansion, although it is not the most divergent term. From eqs. (12.7.15, 12.7.18 and 12.7.19) one obtains

$$f_d(t) \equiv \exp(-i\bar{\varepsilon}_d t)(iDt)^{-g^2} \quad (|t| \rightarrow \infty), \quad (12.7.21)$$

the Fourier transform of which yields the final form of $F(E)$:

$$F(E) \rightarrow \begin{cases} 0 & (E + \varepsilon_c \rightarrow \bar{\varepsilon}_d - 0) \\ (M_d + M_s \Delta_d / V)^2 [D\Gamma(g'^2)]^{-1} [(E + \varepsilon_c - \bar{\varepsilon}_d)/D]^{-(1-g'^2)} & (E + \varepsilon_c \rightarrow \bar{\varepsilon}_d + 0), \end{cases} \quad (12.7.22)$$

where Γ is the Gamma function. It is of the same form as (12.6.5) for the *photoelectron* spectrum, with the difference that the normalized scattering potential $g \equiv -\rho v$ in (12.6.3) is now to be replaced by g' of (12.7.20) which represents the second-order scattering of a conduction electron at ε_F via the intermediate state d.

(A-2) Spectral region: $E \sim \bar{\varepsilon}_F - \varepsilon_c$ where $\bar{\varepsilon}_F = \varepsilon_F + \Delta_g$ and Δ_g is the self-energy of the state $|g\rangle$ due to H' . Here $\bar{\varepsilon}_F - \varepsilon_c$ is the threshold energy for the core electron to be excited into the unoccupied states of the conduction band. In this case it is convenient to rewrite $F(E)$ as

$$F(E) = \left[\pi^{-1} \text{Im} \sum_{k>} \sum_{k'>} G_{kk'} \right] [M_s + VM_d/(E + \varepsilon_c - \varepsilon_d)]^2 \quad (12.7.23)$$

with the use of (12.7.11, 12.7.12). One can calculate $G_{kk'}$ by taking account of the self-energy correction of the final state with a d hole given by the diagram of Fig. 12.10 whose imaginary part is responsible for the lifetime broadening of the Fermi edge mentioned before:

$$\begin{aligned} \Sigma_k(z) &= V^2 \sum_{k'<} (z - \varepsilon_k - \varepsilon_d + \varepsilon_{k'})^{-1} \\ &= \rho V^2 \log[(z - \varepsilon_k - \varepsilon_d + \varepsilon_F)/(z - \varepsilon_k - \varepsilon_d - D + \varepsilon_F)] \\ &\sim \Delta_g + i\pi\rho V^2 \quad (\text{when } E + \varepsilon_c \sim \varepsilon_k \sim \varepsilon_F), \end{aligned} \quad (12.7.24)$$

$$G_{kk}(z) = [z - \varepsilon_k - \Sigma_k(z)]^{-1}. \quad (12.7.25)$$

One obtains a rough description of the lineshape near the Fermi energy from

$$\left[\pi^{-1} \text{Im} \sum_{k>} G_{kk} \right] \sim \rho [(1/2) + (1/\pi) \tan^{-1} \{(E + \varepsilon_c - \varepsilon_F)/\pi \rho V^2\}]. \quad (12.7.26)$$

Namely, the Fermi edge is blurred out by the lifetime width $2\pi\rho V^2$ of the d hole. In an improved approximation the lineshape is given by

$$F(E)/\rho M_s^2 = \pi^{-1} [\text{Im}\{J + \rho V^2 J^2/(z - \varepsilon_d - \rho V^2 J)\}] \\ \times [1 + V M_d/M_s(E + \varepsilon_c - \varepsilon_d)]^2, \quad \text{where } J \equiv \rho^{-1} \sum_{k>} G_{kk}. \quad (12.7.27)$$

The lineshape in the entire region (A-1) through (A-2) is as already shown schematically in Fig. 12.7 whereas the derivation of (12.7.27) and the numerical calculations are presented in Ref. [61].

Case (B) $\varepsilon_d > \varepsilon_F$

(B-1) Spectral region: $E \sim \bar{\varepsilon}_F - \varepsilon_c$. With the use of (12.7.23) one calculates the generating function by taking advantage of the linked cluster expansion as follows:

$$f_S(t) = \int_{-\infty}^{+\infty} dE \exp(-iEt/\hbar) \left[\pi^{-1} \text{Im} \sum_{k>} \sum_{k'>} G_{kk'} \right] \\ = \sum_{k>} \sum_{k'>} \langle g | a_k \exp[-i(H - E_g)t/\hbar] a_{k'}^\dagger | g \rangle \\ = L_2(t) \exp[L_3(t)], \quad (12.7.28)$$

where $L_2(t)$ and $L_3(t)$, respectively, are the sums of all the connected diagrams with and without external lines. In the most divergent term approximation, they are given by the diagrams in Fig. 12.9 and expressed as

$$L_2(t) = (\rho/it)(iDt)^{2g^1} \exp(-i\varepsilon_F t), \quad (12.7.29)$$

$$L_3(t) = -i\Delta_g t - g^{12} \ln(iDt). \quad (12.7.30)$$

Putting eqs. (12.7.29, 12.7.30) into (12.7.28) and performing the Fourier transformation, one finally obtains

$$F(E) \rightarrow \rho [M_s + V M_d/(\bar{\varepsilon}_F - \varepsilon_d)]^2 [\Gamma(1 - 2g^1 + g^{12})]^{-1} \\ \times [(E + \varepsilon_c - \bar{\varepsilon}_F)/D]^{-(2g^1 - g^{12})} \quad (E + \varepsilon_c \rightarrow \bar{\varepsilon}_F + 0), \quad (12.7.31)$$

whereas $F(E) = 0$ when $E + \varepsilon_c < \varepsilon_F$. The singular lineshape is shown schematically in region (B-1) of Fig. 12.8.

It is interesting to note that the threshold singularity of region (A-1) is of the same form as that of the photoelectron spectrum of normal metals, while that of region (B-1) is of the same form as that of the absorption spectrum of normal metals,⁵⁵ except that $g = -\rho v$ is now replaced by g^1 given by (12.7.20). In the

former cases, the core electron is excited to a *definite level*: the pulled-down d level ε_d or the conduction band level corresponding to a definite kinetic energy ε of the photoelectron outside. The singularity can then be described in terms of the sideband due to simultaneous excitation of conduction electrons across the Fermi level ε_F as was done in deriving (12.7.22) as well as (12.6.5). In the latter cases, the core electron is excited to one of the vacant levels above ε_F , being now indistinguishable from other electrons simultaneously excited from below ε_F . The negative power of the singularity in these latter cases is obtained by replacing g' by $1 - g'$ (or g by $1 - g$) as argued by Hopfield,⁴⁸ and hence by replacing $(1 - \alpha') \equiv (1 - g'^2)$ by $[1 - (1 - g')^2] = 2g' - g'^2$.

(B-2) Spectral region: $E \sim \bar{\varepsilon}_d - \varepsilon_c$. The transition of the core electron to the d state interferes with that to the s states through the s-d mixing interaction. It is well known that when the excited states of the optical transition contain a discrete state superposed on a background of continuous states with which it interacts, a characteristic asymmetric peak with a dip on one side (called antiresonance) appears in the absorption spectra, the so-called Fano effect.⁶² One defines the energy shift, Δ , and the level broadening, Γ , of the state $a_d^\dagger|g\rangle$ by the equation

$$G_{dd} = [z - \varepsilon_d - \Delta - i\Gamma]^{-1}. \quad (12.7.32)$$

Putting this into (12.7.14), one can rewrite $F(E)$ in the following form which is the same as Fano's general formula⁶² for the antiresonance:

$$\begin{aligned} F(E) &= (M_s^2 \Gamma / \pi V^2) [E + \varepsilon_c - \varepsilon_d + (V M_d / M_s)]^2 / [(E + \varepsilon_c - \varepsilon_d - \Delta)^2 + \Gamma^2] \\ &= (M_s^2 \Gamma / \pi V^2) (q + \eta)^2 / (1 + \eta^2), \end{aligned} \quad (12.7.33)$$

where

$$\eta = (E + \varepsilon_c - \varepsilon_d - \Delta) / \Gamma, \quad (12.7.34)$$

$$q = [(V M_d / M_s) + \Delta] / \Gamma. \quad (12.7.35)$$

If one assumes that Δ and Γ are slowly varying function of E , the lineshape is determined mainly by the value of q : when $q = 0$, the spectrum shows a dip, i.e., a perfect antiresonance, and as $|q|$ increases an asymmetric peak grows. The asymmetric peak is located on the high- or low-energy side according as $q >$ or < 0 . One can obtain the value of q in the present model by calculating Δ and Γ in an appropriate approximation. See Ref. [61] for the results. The Fano resonance in the case of $q < 0$ is shown schematically in (B-2) region of Fig. 12.8.

Case (C) ε_d changing across ε_F

The negative power of singularities at (A-1) and (B-1), given by $(1 - g'^2)$ and $(2g' - g'^2)$, respectively, tends to ∞ as $|\varepsilon_d - \varepsilon_F| \rightarrow 0$ (see eq. (12.7.20)), indicating

the breakdown of the theory based on the most divergent term approximation. However, according to an improved description in terms of the phase shift of the scattering as mentioned in the latter half of Section 12.6, the power changes continuously from the former expression to the latter as we increase the parameter ε_d across ε_F . (See Ref. [61] for the detailed argument.)

Suzuki *et al.*⁶³ observed the $N_{4,5}$ absorption spectra of La and Ce. They also determined the Fermi level from photoelectron spectroscopy. According to their results, there is no sharp Fermi level threshold at the location of the Fermi level while there exists one or several sharp peaks, which correspond to the electronic transition from 4d to 4f levels, *on the low-energy side* of the Fermi level. These unusual features can be described qualitatively by the case (A) mentioned above.

The emission spectra corresponding to the annihilation of the core hole at absolute zero of temperature can be calculated in the same way. The results are shown schematically in Fig. 12.11. The negative power of the singularity is always equal to that of the absorption spectra in cases (A) through (C) to (B).

The photoelectron spectrum of the same model system was also studied.⁶⁴ The results corresponding to Figs. 12.7 and 12.8 are shown in Fig. 12.12. The physical

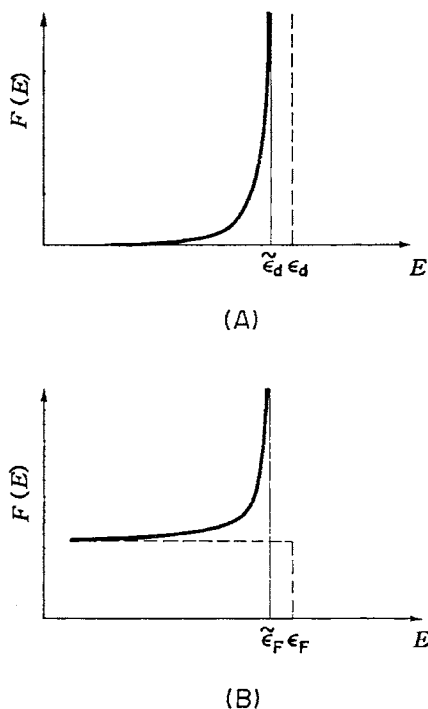


Fig. 12.11 Schematic lineshapes of emission spectra in case (A) $\varepsilon_d < \varepsilon_F$ and case (B) $\varepsilon_d > \varepsilon_F$. The spectra in the absence of s-d mixing are shown with dashed lines.

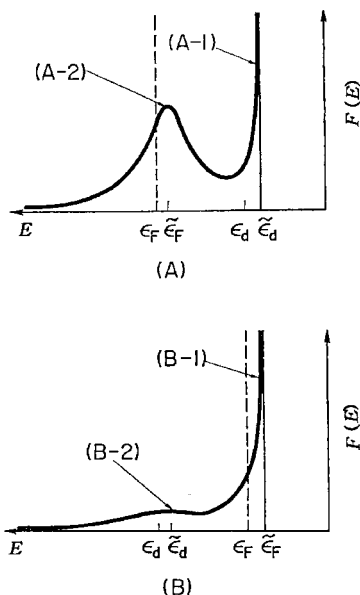


Fig. 12.12 Schematic lineshapes of photoelectron spectra in the case of (A) $\epsilon_d < \epsilon_F$ and (B) $\epsilon_d > \epsilon_F$. The spectra in the absence of s-d mixing are shown with dashed lines.

argument on the results in comparison with the photoabsorption spectra is left for the interested reader.

The two-peak structure as in case (A) of Fig. 12.12 has also been pointed out for insulating compounds of the transition elements, with the conduction band of the metal to be replaced by the valence band of the insulator, and hence with no singularity in the (A-1) component.⁶⁵

The simple model presented above was extended to more realistic ones, taking account of the finiteness of ϵ_d^0 , the orbital and spin degeneracies and hence more electrons in the d shell with their large Coulomb interactions, the exchange interaction between the core hole and the d electrons, spin-orbit coupling and so on. As for the details and the comparison of theoretical results with observations, see the comprehensive review articles by Kotani and collaborators⁶⁶⁻⁶⁸ and the references therein. Of particular interest is the effect of valence mixing reflected in the spectra which involves problems similar to those discussed in this section. The last-referenced review, Ref. [68] deals with resonant inelastic X-ray scattering in general, which is not covered by the present book.

As for the study of site-selective chemical reactions by core-hole excitation (a molecular scalpel, so to speak), mentioned at the beginning of this chapter, see Refs. [69, 70].

13

Photo-induced structural changes

Overview

In this chapter, we will consider the origin and nature of interatomic bonds in the context of optical transitions and subsequent atomic rearrangements. In ordinary textbooks on the solid state, the description of interatomic bonds is put in earlier chapters because that is the very basis of the microscopic structures of matter inclusive of the condensation mechanisms which govern their thermal, electric, magnetic and optical properties. The reason we have delayed our consideration of these aspects until the penultimate chapter in this book is that optical excitation often causes a drastic change in the interatomic bond (inclusive of bond creation and annihilation) which is reflected in and thus revealed by, the optical spectra (emission, as well as absorption spectra) themselves. Thus, the spectroscopic study sheds new light on the nature of interatomic bonds in condensed matter.

Of course, we are well aware of the nature of interatomic bonds in familiar materials. However, interatomic forces acting upon any particular atom are *altogether in balance in the ground electronic state* of the material (however large an individual force may be), so that there is *no way of singling out a force* between any particular pair of atoms as long as the system stays in the ground state. The best way to detect the individual interatomic bonds is to change them locally by optical excitation to an excited electronic state. The imparted imbalance of the interatomic forces will cause the atoms to displace from their Franck–Condon (FC) configuration to a new equilibrium configuration, the relaxed excited state (RES), as shown by the C.C. (configuration coordinate) model of Fig. 13.1, where the abscissa representing the interaction mode Q is drawn *aslant* (instead of horizontal) for the purpose of using this C.C. model in a different situation to be mentioned later. The optical emission spectra from this RES, if any, (not shown in this figure, see instead the down arrows in Figs. 4.1 and 4.4) and its correlation with the absorption spectra (upward arrow) will further reveal the nature of this relaxation and hence of the relevant interatomic

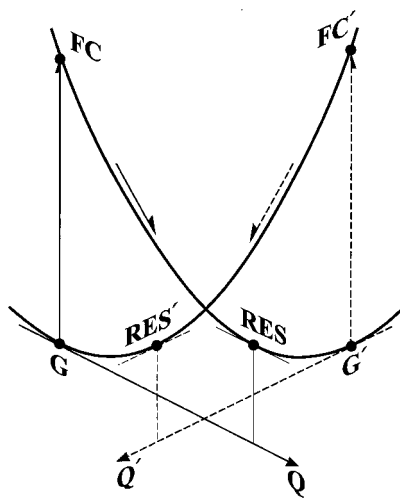


Fig. 13.1 Bilaterally constructed configuration coordinate model useful for nearly bistable situation, with different abscissas GQ and $G'Q'$ for different situations.

bonds, as was described in Section 10.6. Thus, the nature of the interatomic bonds and the mechanism of condensation *in the ground state can be revealed only by the study of excited states* in which the destroyed balance of forces is healed (by the relaxation) and finally reconstructed (in the RES)¹.

Underlying this viewpoint is a *latent possibility* that the lowest point of the excited electronic state could be lower than that of the ground state in a different situation with somewhat different values of the physical parameters, as shown again by Fig. 13.1 with the same pair of adiabatic potentials but with a new abscissa $G'Q'$ slanted oppositely to the abscissa GQ (symmetrical drawing is not essential). Namely, what was considered as the excited electronic state is now the ground electronic state in this new situation.

The purpose of the present chapter is to study photo-induced structural changes, in particular those arising in nearly bistable situations as depicted in Fig. 13.1. In the case of bulk insulating crystals, the lowest excited state in the FC configuration is a free exciton, which becomes self-trapped (RES) in the deformable lattice when the exciton–phonon coupling is strong enough. Namely, the relaxation consists of electronic localization and atomic displacements. If the energy of a self-trapped exciton (STE) were lower than that of the ground state, STEs would be spontaneously generated (without optical excitation) at *all sites* of the entire lattice, forming an electronically as well as structurally new phase of the same matter. The local excitation in this new phase could be described with reference to the new abscissa $G'Q'$, whilst still retaining the same pair of the electronic states as before since the excited electronic state in the old phase is the ground electronic state in the new phase and

vice versa. Which is the real ground state and which is the false ground state is a subtle problem in a nearly *bistable* situation. Although such a situation seems to be realized rather seldom in nature, it is worthy of serious consideration in order to deepen our understanding of electronic–structural phase transitions. In the case of a bulk phase transition, the new and old abscissas, $G'Q'$ and GQ , respectively, should be macroscopic *order parameters* obtained by an appropriately normalized linear combination of the interaction modes. Examples of the order parameter will be seen in Section 13.4 (the dimerization parameter q) and Section 13.6 (alternating mode $N^{-1/2}q_a$ given in eq. (13.6.19)). Also, the interactions among STEs should be included in the new adiabatic potentials.

As an obvious Gedankenexperiment for a false ground state (G') which is far from bistability, let us consider an assemblage of the same number of *neutral* alkali and *neutral* halogen atoms arrayed alternately on a three-dimensional lattice with a sufficiently large lattice constant r with weak interatomic forces of the van der Waals type, and then excite a valence electron of an alkali atom optically to the lowest unoccupied orbital of a neighboring halogen atom at distance $r(FC')$. The excitation energy is positive if the redox potential (defined here as ionization energy I of alkali atom minus electron affinity A of halogen atom) plus the Coulomb potential ($-e^2/4\pi\epsilon_0 r$) of the thus-produced ion pair is positive, namely if the lattice constant r is sufficiently large. The excited state would then relax so as to decrease the nearest neighbor distance r and so make this ionized pair more stable than the neutral pair state. Since this sign reversal takes place already at r significantly larger than a typical lattice constant, the ionized pair would spontaneously be generated at every site so as to form a condensed phase of ionic pairs, namely, an ionic crystal which we know well. This is the obvious reason why an alkali halide forms an ionic crystal rather than a van der Waals crystal.

Although this is a trivial example, one can thereby realize that to imagine the false situation on the opposite side of the reality is sometimes useful for a deeper understanding of the real ground state and the nature of bonding, and that the optical absorption and emission spectra provide a powerful means of detecting the nature of this bond from the opposite side. This chapter contains many examples of situations in which the relative stability of the real and false ground states is a more subtle, and hence more instructive, problem. Some systems are subject to a structural phase transition, which need not be thermally induced but can be photo-induced, revealing intriguing dynamics of the growth and decay of the metastable domain.

Before proceeding to these examples of *bulk* phase transitions, however, we will present well-established examples of the covalent bond and the local photochemical reaction in the following two sections.

13.1 Role of p-holes in formation of covalent bonds

As is well known and was also mentioned in Section 12.1, the one-electron energy in an atom depends mainly on the principal quantum number n (corresponding to a *shell*) and to a lesser degree on the azimuthal quantum number ℓ (corresponding to a *subshell*) in low-mass elements. However, the large centrifugal potential for $\ell \geq 2$ gives rise to an inverted order such that, energetically, $4s < 3d (< 4p)$, $5s < 4d (< 5p)$, $6s < 5d$, $4f (< 6p)$, etc. These d and f orbitals are confined to the inner region of atoms compared to s (and p) orbitals with higher quantum number n in spite of being energetically higher than the latter (for s but not for p). As a result, d and f subshells which are incompletely filled (due to strong Coulomb repulsion against a further electron entering the same orbit) contribute to magnetic properties with their resultant spin coming from those of the individual d (or f) holes (unoccupied orbitals of the d or f subshell). The incompletely filled p subshell plays the main role in forming interatomic covalent bonds with their directions being those of the p orbitals of individual *holes*. Note that the unoccupied antibonding orbital (rather than the occupied bonding orbital) contributes in effect to net bonding, as will be made more clear in the following argument. The existence of a *hole* (rather than an *electron*) is essential in both cases; a complete subshell contributes neither to the intra-atomic magnetic moment nor to the interatomic covalent bond.

Let us consider how the exchange (transfer) of a hole(s) in the outermost p subshell between two atoms gives rise to the interatomic bond. For this purpose we start with a column 0 element with a closed np subshell (a rare gas with $n \geq 2$) which is known to be chemically inactive, and shift leftward to elements with smaller Z , with ν holes in the p subshell ($0 < \nu \leq 3$ where 3 is the maximum number of linearly independent p *orbitals*) in order to study the possible number of interatomic bonds with like atoms, by considering the well-known *empirical* facts which formed the basis of the periodic table.

A halogen atom with $\nu = 1$ (column VII) can form *one* covalent bond with another identical atom but not more, resulting in diatomic molecules such as F_2 , Cl_2 , Br_2 and I_2 . While there are *two* holes in a pair of halogen molecules in this example, there is evidence that even *one* hole in a pair of rare-gas atoms is enough to form a molecule. Consider an optically-excited atom in a rare gas with an electron in a higher state and a hole in the np subshell. The hole has the effect of attracting another atom still in the ground state through its resonance transfer between the two atoms, resulting in the formation of an *excimer* (a molecule in an excited state) as shown schematically in Fig. 13.2. The excited electron plays no direct role in the excimer formation except that of leaving a hole behind. As soon as the electron is optically de-excited (p-hole annihilated), the interatomic attractive force disappears and the two atoms become separated as before. The photon thereby emitted can

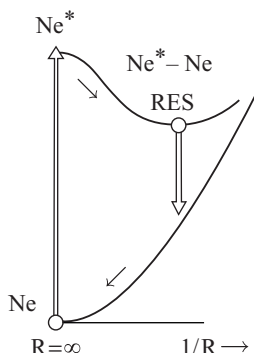


Fig. 13.2 Configuration coordinate model for an optically-excited rare-gas atom to form an excimer $\text{Ne}^* + \text{Ne} \rightarrow \text{Ne}_2^*$.

be used as a laser light source because of population inversion at this interatomic configuration (optical pumping produces RES while there is no population in the Franck–Condon de-excited state). This is nothing other than the excimer laser.

As another example of *one* *p*-hole forming a covalent bond, consider the optical interband excitation of an alkali-halide crystal which produces a positive hole in the valence band consisting of np orbitals of halide ions X^- . A hole naturally wants to be in the highest possible electronic state, which is an *antibonding* state between a nearest pair of neighboring X^- ions. The attractive force mediated by the positive hole makes the two X^- ions come closer together. Namely, an X_2^- molecule, or self-trapped hole of two-center type, arises in the crystal, which plays a central role in a variety of photochemical processes, as will be described in Section 13.2.

Proceeding to the next example, let us consider a chalcogen atom with $\nu = 2$ (column VI). It can form two covalent bonds with two neighboring atoms of the identical species, with an angle somewhat greater than 90° , as can be seen from examples such as the bent ozone molecule O_3 , the crown-shaped S_8 or Se_8 molecule consisting of a closed ring of a zig-zag chain (which can further aggregate to form a crystal) and zig-zag chains of Se and Te arrayed regularly to form a three-dimensional crystal or disordered to form the amorphous state.

The situation is more varied in $\nu = 3$ elements (column V); besides the 3-coordinated structures such as the tetrahedral P_4 molecule and molecules consisting of N, P, As or Sb with three neighbors of H or a halogen, they can also form simple molecules such as N_2 , NO and various compounds with apparently more than three neighbors which may be ascribed to the lone-pair electrons (electrons *not* participating in the *covalent* bond) or to the removal of electrons to neighbors (ionic bonds).

The situation is even more varied with elements with $\nu = 4$ (column IV) where three independent np orbitals are not enough to make holes. In order to have four

covalent bonds with four neighbors, one has to resort, at the cost (from the hole viewpoint) of intra-atomic energy, to a lower-lying ns orbital as another independent basis for hole orbitals. In other words, an environment of tetrahedral symmetry will make the well-known set of four sp hybrid orbitals

$$\begin{aligned}\phi_1 &= 2^{-1}(s + p_x + p_y + p_z), & \phi_2 &= 2^{-1}(s + p_x + p_y - p_z), \\ \phi_3 &= 2^{-1}(s + p_x - p_y - p_z), & \phi_4 &= 2^{-1}(s - p_x - p_y - p_z)\end{aligned}\quad (13.1.1)$$

become energetically degenerate so as to form four covalent bonds with tetrahedrally coordinated neighbors, as is seen in the covalent crystals of C (insulator), Si and Ge (semiconductors). This diamond structure is so stable that even the column V and column III elements can be combined to form compound semiconductors with the similarly framed zinc-blende structure by transferring an electron from the column V to the column III element. The valence and conduction bands consist of bonding and antibonding orbitals as mentioned later, so that the two statements of (i) *electrons* occupying the stable *bonding* orbitals and (ii) *holes* occupying the unstable *antibonding* orbitals are equivalent here.

However, the number-of-holes picture (hereafter the hole picture) for covalent bonds starting from a column 0 element and moving to the left along a row of the periodic table is obviously more natural than the number-of-electrons picture (hereafter the electron picture) starting from the opposite end of the sp^3 shell, since the s orbital with which we have to start has no degeneracy and hence no directionality, being irrelevant to the notion of an interatomic bond. In fact, column I as well as column II elements hardly ever form molecules but coagulate together to form metallic crystals in which the s electrons are quite mobile, being shared by many atoms – the so-called *metallic bond*. The asymmetry of the hole and electron pictures within the sp^3 community comes also from *one* s orbital versus *three* p orbitals with different energies. The electron picture within the $(np)^3$ community is not so useful as the hole picture, presumably because of the strong influence from ns below np , especially for $n = 2$. For heavier elements with greater ns – np splitting, however, there are the interesting examples of $PbCl_2$ and PbI_2 crystals in which an electron in the conduction band consisting mainly of the $6p$ orbitals of Pb is self-trapped symmetrically around a pair of Pb^{2-} ions. These two compounds are examples of the relatively few cases of *electron* self-trapping (see the end of Section 9.6 inclusive of its references).

The column III elements do not really fit into either picture, presumably because the highest occupied level is just above the s – p borderline. In fact, the lightest element in this column, B, in which the s – p splitting is smallest, shows such unique and versatile chemical bonds that it is even nicknamed “a lone wolf” among the elements of the periodic table.

An exception to the asymmetry of the hole and electron pictures is the first row of the periodic table. Due to the absence of the 1p state, (1s)² already forms a closed shell. Hydrogen can be viewed either as a one-electron system or as a one-hole system. In fact, the existence of a diatomic molecule H₂ as well as the formation of metallic hydrogen under extremely high pressure are consistent with the amphoteric picture.

Let us now give a simple argument for the preference of the hole picture for the covalent bond in the context of the Heitler–London-like theory. Instead of the 1s atomic orbital, however, we begin with the 2p orbital without 2s mixture. Consider two identical atoms with at least one 2p electron, approaching each other along the *z*-axis. One can then imagine a *bonding orbital*

$$\sigma_b = 2^{-1/2}(2p_{z1} - 2p_{z2}) \quad (13.1.2)$$

formed between the atoms 1 and 2 (the $-$ sign gives a molecular orbital of even parity because of the odd atomic orbital *p*). Considering the angular parts of the two atomic orbitals and the absence of a node in their 2p radial wave function, one finds that the orbital function σ_b is strongly concentrated in the interatomic region where the attractive potentials due to the two atoms have a simultaneously large negative contribution to the total energy (provided the interatomic distance is appropriate). This is the well-known reason why the bonding orbital has one-electron energy lower than an atomic orbital, and at the same time justifies the condition stated within the parenthesis just above since the interatomic distance is so determined as to maximize the negative contribution mentioned there. The same argument applies to shells with $n > 2$; although the radial part of the *np* orbital has nodes and the behavior of the bonding orbital in the interatomic region is more complicated than in the 2p shell, the optimum interatomic distance is so determined as to maximize the negative energy of the bonding state.

In contrast, the antibonding orbital σ_a , given by (13.1.2) with the $-$ sign replaced by $+$, has almost no energy gain since it avoids this comfortable interatomic region. Thus, the antibonding state has an energy much larger than the bonding state.

If the two atoms have a closed-shell structure with *np* orbitals completely occupied, we have to fill up the bonding and antibonding orbitals each by two electrons with up and down spins in addition to other atomic orbitals np_{xi} and np_{yi} ($i = 1, 2$) which do not participate in the covalent bond. The Slater determinant formed by 12 *np* electrons inclusive of bonding and antibonding orbitals and of up and down spins is nothing other than that formed by 12 atomic orbitals forming closed shells. Summation of the energies of the bonding and antibonding orbitals should result in no net contribution to the interatomic binding energy. The contribution to the binding energy comes only from *holes* in the atomic *np* shell; one or two hole(s) should occupy the antibonding orbital with higher energy than the bonding orbital,

thus contributing a negative energy and hence increasing the binding energy as compared with a pair of closed-shell atoms. This is the reason why the hole picture is more natural when one starts with a closed-shell structure. The second and further holes in the atomic np shell, if any, must form covalent bond(s) with other atom(s) in other direction(s) (x and/or y), or stay unbonded. The hole picture described here is equivalent to what is called the $8 - N$ rule.

It is interesting to note that the interbond angles in the elements with $v = 2$ or 3 mentioned above are known to be H_2S : 92° , H_2O : 104.5° , O_3 : 117° , $\text{S-S-S} \dots$: 108° , $\text{Se-Se-Se} \dots$: $103\text{--}106^\circ$, NH_3 : 107° , etc. The deviation of the angle from 90° expected for pure np_x , np_y , \dots orbitals indicates the corresponding mixing of the ns orbital at the cost of np – ns splitting energy so as to counterbalance the repulsive energy between the ligand atoms of the covalent bonds.

13.2 Photo-chemical processes in alkali halides

Alkali halides are ideal ionic crystals consisting of cations and anions both with closed-shell structures, bound together by large Coulomb and small van der Waals forces counterbalanced by short-range repulsive forces. Their electronic structure is very simple, with a valence band consisting of the np subshell of the halide ion (X^-) and a conduction band consisting of the n 's and possibly the $(n' - 1)d$ ($n' \geq 3$) subshell of the alkali ion (M^+). To a first approximation, the binding energy consists of the electrostatic Madelung potential minus the redox potential (the ionization energy of the alkali atom minus the electron affinity of the halogen atom).

In contrast to this simple ground state, the electronic excited states behave quite variably, especially in their lattice-relaxed configurations. The lowest optical excitation is to bring back an electron from an anion to a cation which amounts to *reduction* in chemical terminology. The sudden disappearance of the electronic charges on the relevant ions and hence of the Coulomb energy which was responsible for a significant part of the binding energy, as mentioned above, naturally causes a drastic change in the balance of the interatomic forces, resulting in large lattice relaxations with various patterns, as already mentioned in Section 9.6. The lattice relaxation can be so large that defect pairs of an anion vacancy with an electron (F-center) and an interstitial anion with a positive hole (H-center), or their precursor (*off*-center-type self-trapped exciton (STE)) are produced thereby (see Fig. 13.3, (d) + (d') or (c)), according to experimental studies. A single positive hole, produced under ionizing radiation with energy greater than the bandgap, becomes self-trapped symmetrically around a pair of anions which are thereby brought closer together than their normal distance, forming a V_K -center with net charge $+e$ as shown in Fig. 13.3(a), which is a typical pattern of a self-trapped hole (STH) in ionic crystals. This STH can again capture an electron around its attractive potential, forming an *on*-center-type STE as shown in Fig. 13.3 (b). Throughout

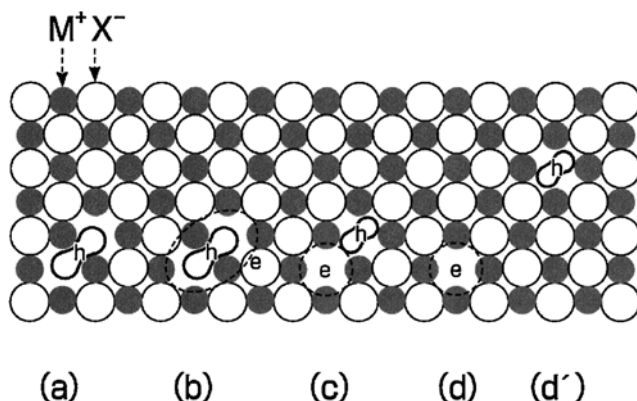


Fig. 13.3 Various defects formed in a photoexcited alkali-halide crystal. Reproduced from the article by Song and Kayanuma.²

(a)–(c) and (d'), one finds that a positive hole prefers to be self-trapped around a pair of anions by bringing them closer than their normal distance in the host crystal *irrespective* of the different environments: in the intercation (a, b) or interstitial (c, d') positions, under the absence (a) or presence (b, c, d + d') of a nearby electron, under the absence (a, b) or presence (c, d + d') of a nearby anion vacancy. It indicates the *primary* importance of the hole-mediated interhalide bond mentioned in Section 13.1 within the excited-state subspace with an electron and a hole.

At this point, we have to touch upon some of the historical turns and twists leading finally to the established identifications of experimentally-found centers with defect models shown in Fig. 13.3. After the model of STH shown in (a) was established on the basis of an electron spin resonance study (see Ref. [21] of Chapter 9), it was long believed that an STE is an electron captured around an STH keeping its parity (inversion symmetry) (b). Although the broken parity was suspected through an electron–nuclear double resonance study,³ it was not until the comprehensive theoretical study by Song and collaborators⁴ (see also Ref. [53] of Chapter 9) that certain alkali halides were found to have off-center structures (as shown by (c)) as the most stable configuration of the STE. This meant that what had so far been considered to be a pair of F- and H-centers should be reclassified into the off-centered STE (the nearest pair, (c)) and more distant pairs (d + d'). Kan'no and collaborators (see Ref. [63] of Chapter 9) clarified that there are three types of stable (or metastable) configurations of STE, on-center type I (Fig. 13.3(b)), weakly off-center type II (not shown in Fig. 13.3) and strongly off-center type III (Fig. 13.3 (c)) depending on the alkali halides, by elaborate comparative study of their emission spectra as already mentioned in Section 9.6.

Starting with an ionized electron–hole pair produced by radiation, the first relaxation process to arise is breaking of translational symmetry, resulting in the *on-center* STE (b). The next to arise would be breaking of parity, resulting in an

off-center self-trapped exciton (c) or a pair of an F-center and H-center ($d + d'$) as if the X_2^+ molecular ion is repelled by an electron of opposite charge. How does this repulsion arise against the Coulomb attraction?

Kondo, Hirai and Ueta,⁵ through their transient spectroscopic study of the formation process of F- and H-centers by irradiation with a pulsed electron beam, clarified that the centers are produced within nanoseconds, which is much shorter than the radiative lifetime of the self-trapped excitons, and concluded that the formation takes place while the electron is in the excited Coulomb-captured state(s) around the STH. Independently, Itoh and Saidoh⁶ considered the *down-hill* adiabatic potential of the STE towards the *parity-breaking* displacement Q_2 of X_2^- as a whole along its molecular axis (along (110) in the plane Fig. 13.3, which is defined here as the z -direction), in order to explain the formation of F- and H-centers, $(b) \rightarrow (c \text{ or } d + d')$. Toyozawa⁷ reconciled the two pictures by ascribing the down-hill potential to the pseudo Jahn-Teller effect of the ground ($\sim 1s$) and the excited ($2p_z$) states of the electron captured by the STH. The pseudo J-T effect was assumed to be significant *only at* the interanion distance where the $2p_z$ and $1s$ levels are close enough (so that also the nonradiative transition $2p_z \rightarrow 1s$ readily takes place) as shown schematically in Fig. 13.4, in order to keep two different channels, one leading to an STE and another to an F-H pair which were considered to be different entities at that time. However, Song's identification of the STE with the nearest F-H pair

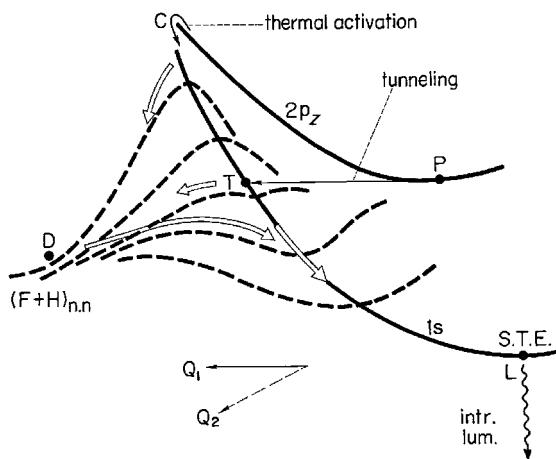


Fig. 13.4 Configuration coordinate model with two competing channels of relaxation, one leading to L, the self-trapped $1s$ exciton, from where intrinsic luminescence takes place, and another to D, where a nearest pair of F- and H-centers are formed. P is the relaxed state of the $2p_z$ exciton, from which atoms can tunnel quantum-mechanically to the point T on the adiabatic potential of the $1s$ exciton. C is the crossing point of the $2p_z$ and $1s$ -exciton adiabatic potentials where thermally activated de-excitation takes place.⁷

(Fig. 13.3(c)) resolved the confusion through their calculation according to which the down-hill potential due to the pseudo J-T effect *exists even for the equilibrium interanion distance* in the V_K -center,⁸ namely under a $1s-2p_z$ difference as large as ~ 1 eV. In alkali halides with a type I STE, however, the model of Fig. 13.4 may be applicable.

A simpler way of explaining the off-center stabilization ((b) \rightarrow (c) of Fig. 13.3) is that the electron prefers being captured by charge $+e$ of an *entire* anion vacancy as in (c) to being captured by two charges each with $+e/2$ due to *half*-anion vacancies as in (b). The molecular ion X_2^- separating them acts to eliminate the resonance effect of the electron between them.

However one views it, it is the much lighter electron which drives the heavy X_2^- molecular ion in the (110) direction. Namely, a pair of an electron and a hole, particles with effective masses of the order of m , created by a photon with energy $\sim \varepsilon_g$, produce a pair of defects with much larger effective masses of the order of anion mass M . The hole-mediated interanion bond causes breaking of the translational symmetry of a free exciton, while the adiabatic down-hill potential due mainly to the electron-ion Coulomb interactions provides X_2^- with momentum, breaking its parity.

13.3 Lattice relaxation of exciton across bandgap and its radiative and nonradiative annihilation

We will now study various possible cases of lattice relaxation of the photoexcited state, such as were described in the preceding section, in the context of competition between two channels of de-excitation, radiative and nonradiative, by extending the configuration coordinate (C.C.) Q of Fig. 10.10 to the point where the adiabatic potentials of the ground and excited states intersect each other.⁹ Consider the exciton band with width $2B$ and bottom energy at E_0 measured from the ground state, and local lattice distortion Q which gives rise to the local attractive potential $-cQ$ ($c > 0$) for the exciton. E_0 will hereafter be called the *exciton gap*. The energy of a localized exciton will tend to the asymptote $E_c(Q) = E_0 + B - cQ$ as shown by the solid line in Fig. 13.5(a), since complete localization gives rise to kinetic energy B according to the uncertainty principle as mentioned in Section 10.6. Assuming for simplicity the linearity of the exciton-phonon interaction throughout the exciton gap, one finds that the energy $E_c(Q)$ of the localized exciton intersects the ground-state energy at $Q_0 = (E_0 + B)/c$. What is the physical meaning of Q_0 ? In insulators, the exciton gap E_0 , the interatomic charge transfer and the lattice structure are self-consistently correlated with one another as stated in the Overview. This means that the intersection Q_0 is so large as to cause an *entire* change

$E_R \equiv W_e(0) - W_e(Q_m) = c^2/2$. Note that $E_0 \geq W_e(Q_m)$ according as the coupling constant $g \equiv E_R/B \leq 1$ (only the latter case is shown in Fig. 13.5(b)). Note that we have not specified the spatial pattern of atomic displacements represented by the “local” mode Q . Obviously, a single material can have more than one type of STE with different modes (e.g., with different symmetries) and different E_R . The argument below applies to any of these cases.

Let us now study how the adiabatic potential of the excited state, $W_e(Q)$, intersects with that of the ground state, $W_g(Q) = E_L(Q)$. Four typical cases, α , β , γ and U are shown in Fig. 13.5(b) in order of *decreasing* exciton gap E_0 with fixed values of B and $E_R(>B)$. Case S_α is characterized by the *inequality* that the adiabatic potential at the intersection, given by $W_e(Q_0) = W_g(Q_0) = (E_0 + B)^2/2c^2$, is greater than the energy E_0 of the free exciton at the band bottom wherefrom the exciton started to relax. According to the Dexter–Klick–Russel criterion¹⁰ originally proposed for the presence–absence argument of the color-center luminescence based on the assumption that the *nonradiative* transition can take place *only at the intersection* of two adiabatic potentials (see Refs. [11–13] for the quantum theory of nonradiative transition at the intersection), the above-mentioned *inequality* excludes the possibility of “transient” nonradiative annihilation (during the relaxation) of the exciton, because a “thermally-activated” nonradiative process is improbable insomuch as $k_B T$ is small compared to $W_e(Q_0) - W_e(Q_m)$. This is a situation favorable for the *high quantum yield* of F and S luminescences as well as for the F–S thermal equilibration. In Fig. 13.6 representing the (E_0, E_R) -plane with a fixed value of B , case α is realized in the region S_α which is bounded by the above inequality: $E_R < (E_0 + B)^2/4E_0$ and the inequality: $E_R > B$ for self-trapping.

In the opposite situation, $W_e(Q_0) = W_g(Q_0) < E_0$, the transient nonradiative transition of the exciton can take place rather efficiently at Q_0 within one or a few

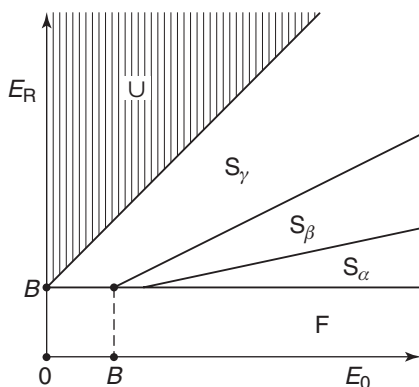


Fig. 13.6 Classification of relaxed excited states.⁹

cycles of damping oscillation of the lattice. The remaining excitons, which may be a small fraction, will finally relax to the self-trapped state Q_m . We must then classify the situation into cases β and γ according as $W_e(Q_m) \geq W_g(Q_m)$ (which is equivalent to $Q_0 \geq Q_m$ and to $E_R \leq (E_0 + B)/2$). In case β , the STE will *also be annihilated radiatively*, emitting a photon with energy $W_e(Q_m) - W_g(Q_m) > 0$. However, the *quantum yield of luminescence is significantly smaller* than in case α . The population of F and S excitons, which are responsible for F and S emission bands, respectively, are *far from thermal equilibrium, being unfavorable to the latter* because of the transient nonradiative sink at Q_0 which is on the S side of the F–S potential barrier. It should be remembered here that the Dexter–Klick–Russel criterion is based on the assumption that Q makes undamped oscillations starting from the unrelaxed point $Q = 0$. In fact, Q is a sort of the interaction mode (see Section 4.4) which, being a superposition of normal modes with different frequencies, makes damped oscillations around the equilibrium point Q_m . According to the dynamical theory of defect reactions due to Sumi,¹⁴ this effect shifts the α – β boundary of Fig. 13.6 upward.

In case γ , the relaxed excited state S_γ can *no longer luminesce* since $W_e(Q_m) < W_g(Q_m)$. Moreover, $Q_m(> Q_0)$ is probably so large that a sort of Frenkel defect is already formed, as was mentioned for Q_0 before. (This may happen even in case β provided Q_0 is close to Q_m ; it is only in a qualitative sense that Q_0 represents the defect-formed state.) S_γ can be *de-excited* into the ground state *only nonradiatively* – by thermal activation over Q_0 at high temperature or tunneling through Q_0 at low temperature – whereby the electron–hole pair and the Frenkel defect are *simultaneously annihilated*.

In Fig. 13.6, the region S_γ should be bounded by another inequality: $W_e(Q_m) = E_0 + B - E_R > 0$. Otherwise (the hatched region U in Fig. 13.6), the relaxed exciton with negative energy $W_e(Q_m)$ (*unstable* exciton, so to speak) would be generated spontaneously (without optical excitation) at every lattice site, resulting in an electronic and structural phase change of the bulk crystal. This means that the *crystal lattice* and its *ground electronic state* from which we started is *not a really stable phase*, corresponding to the situation of Fig. 13.1 with abscissa $G'Q'$. To be more exact, the interactions among the relaxed excitons must be included when they are spontaneously generated, so that the boundary between S_γ and U regions should be shifted to either side as was also mentioned in the Overview. Be that as it may, it is instructive and illuminating to study the stable phase of a many electron–many-nuclei system from the viewpoint of exciton–phonon instability in a hypothetical crystal. We will give some examples in later sections.

Our argument is based on the conceivably simplest model – a *single interaction mode* Q which is *unspecified*, the *linear interaction* of an exciton with a *harmonic lattice* extrapolated as far as *to the configuration of defect formation*. Nevertheless,

the conclusion therefrom qualitatively reproduces the general trend of the experimental results on radiative and nonradiative decay processes of excitons depending upon the relative magnitudes of three material constants: E_0 , B and E_R . A few examples will be mentioned.

Consider the excitons in crystals of aromatic molecules of Frenkel type, where E_0 corresponds approximately to the *intramolecular* excitation energy. For these *soft* aggregates of *hard* molecules, B originates from the intermolecular resonance interactions (consisting of dipole–dipole and exchange interactions, see Section 8.1) and E_R originates from their change with intermolecular configuration, both being of the same order of magnitude but usually much smaller than E_0 . Therefore, excitons in these materials will be located not far from (but on both sides of) the S_α –F boundary. This explains why some aromatic crystals have excitonic luminescence of high quantum yield consisting of F and S components in *quasi-thermal equilibrium* with their relative stability being *readily inverted* by applying pressure.¹⁵

Among wide-gap narrow-band ($E_0 \gg B$) ionic crystals, alkali halides are especially rich in the possible configurations (different modes) of the relaxed excited state – STEs of type I, II, III and F–H pair (see Sections 9.6, 13.2 and Ref. [16]). The STEs belong to S_β in view of the *low quantum efficiency* of their luminescence and their being *far from thermal equilibrium with the free excitons* such that the height of the F to S potential barrier, instead of the F–S energy difference, can be measured in the transient study of luminescence. However, the F–H pair in alkali halides belong to S_γ in view of the *absence of luminescence*. It should be noted in passing that α , β , γ and U in Fig. 13.5(b) are presented for a *variable* bandgap with *fixed* values of B and E_R . The S_β and S_γ in alkali halides represent the existence of different relaxed states with moderate and large E_R and Q_m for fixed E_0 .

A similar situation is found in *amorphous* chalcogens and chalcogenides.¹⁷ For bandgap photoexcitation of a-Se, a-As₂S₃ and a-As₂S₃,¹⁸ there appears a photoluminescence band in the midgap region, which belongs to S_β of our classification. Note that the photoexcitation around the so-called bandgap in amorphous semiconductors includes excitons and band-tail states without being resolved spectrally. Bandgap photoexcitation also causes a structural change as observed by the *photodarkening* of thin films,¹⁹ the red shift of the absorption edge in the Urbach region, which can be partly recovered by subsequent annealing near the glass-transition temperature.²⁰ (It should be mentioned that the Urbach rule mentioned in Section 10.6 for crystalline materials holds also in many amorphous semiconductors where the origin of the tail is to be ascribed partly to distributed lattice defects inherent to amorphous systems and partly to lattice vibrations.) The spectral dependence of photodarkening efficiency has been studied, indicating that it rises at the absorption band edge and saturates or falls beyond that depending on the material.²¹ The

photo-induced defects are to be classified as S_γ because they do not luminesce. In fact, various C.C. models similar to Fig. 13.5(b) have been used independently²² to explain their thermal- and photo-bleaching effects. Since amorphous materials are in one of innumerable metastable states with various types of defects, photostructural change may bring the system to either higher or lower states than the initial state (see Fig. 13.1). We are, so to speak, in the vicinity of the S_γ -U borderline which itself is already blurred as long as we stay in the amorphous, instead of the crystalline, state. It is to be noted that recent *in-situ* observation by EXAFS (see Section 12.2) of the photoexcitation effect of a-Se revealed²³ that the coordination number of Se, normally two, increases slightly (only during the photoexcitation) indicating the formation of a small number of interchain bonds (a pair of atoms with coordination number 3) possibly through the optical excitation of a lone-pair electron towards a neighboring chain. While this local change of the electronic bond disappears as soon as photoexcitation is stopped, the photo-induced structural disorder (mean square relative displacement) remains, contributing to photodarkening, etc.

In the opposite limit of narrow-gap wide-valence-band semiconductors ($E_0 \ll B$ where $2B$ can be taken to represent the valence-band width), one can readily find, in Fig. 13.5, the reason why no STE with luminescence (neither S_α nor S_β) is found in these materials. Even the S_γ region for the nonradiative defect state is much smaller than the F region. In principle, however, one can expect the existence of materials belonging to the S_γ region which are so sensitive to light that nonluminescent and hence long-lived defects are readily formed and stay. Some narrow-gap materials, crystalline or amorphous, may possibly be such. It is to be noted in passing that the glassy state is not in thermal equilibrium.

It is also to be noted that the configuration coordinate model of Fig. 13.5(b) and the phase diagram of Fig. 13.6 can be applied also to molecules by putting $B = 0$. There are many examples of photo-isomerized states of organic molecules.

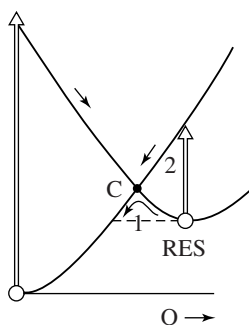


Fig. 13.7 Configuration coordinate model for photo-isomerization of azobenzene which can be reversed thermally (1) or by infrared irradiation (2). Due to Kumar and Neckers.²⁴

In azobenzene,²⁴ for example, the *trans* and *cis* forms are the ground state and RES, respectively, which can be converted to each other by a photon of appropriate energy as schematically shown in Fig. 13.7. As another example of a photochemical reaction, we have already shown the C.C. model for excimer formation in a rare gas in Fig. 13.2. Since the excimer emits a photon with high efficiency, it belongs to the S_α case.

13.4 Neutral–ionic phase transition in quasi-one-dimensional charge transfer compounds

We shall be concerned in this section with charge transfer compounds consisting of donor (D) and acceptor (A) molecules whose redox potential Δ ($\Delta \equiv I - A$ where I is the ionization energy of the donor and A is the electron affinity of the acceptor) is not very different from the Madelung potential gain M in the ionic phase consisting of the ionized donor and acceptor. Such compounds are near the boundary between the neutral (N-) phase and ionic (I-)phase; they are in the N-phase (when $\Delta - M > 0$) or the I-phase (when $\Delta - M < 0$) depending upon the combination of D and A molecules. They are typical examples of the nearly bistable situation depicted by Fig. 13.1. Some compounds close to the boundary ($\Delta - M \sim 0$) take N- and I-phases depending upon the temperature or pressure which cause a small change in the lattice constant a and hence in $M \propto a^{-1}$. In contrast to typical ionic crystals such as alkali halides which are far from the boundary ($\Delta - M$ is far on the negative side), these compounds will provide us with a very good chance of studying spectroscopically how various electronic and structural properties are different *on both sides of*, and change *through*, the neutral–ionic phase transition.

The charge transfer compounds with which we are concerned here are quasi-one-dimensional organic compounds consisting of chains of mixed-stack D and A molecules. Contrary to inorganic ionic crystals, the neutral molecules D^0 and A^0 have an electronically closed-shell structure while the molecular ions D^+ and A^- have unpaired electrons. Among such compounds,²⁵ we mainly consider TTF (tetrathiafulvalene, donor)-CA (p-chloranil, acceptor) which is subject to an N–I phase transition as the temperature is lowered below $T_c = 81$ K or under an applied pressure above 10 kbar at room temperature. The degree ρ of ionicity in the condensed phase, defined by D^0A^0 (isolated) $\rightarrow D^{+\rho}A^{-\rho}$ (condensed), changes abruptly from $\rho \sim 0.3$ to $\rho \sim 0.7$ at the N–I transition, as estimated from a change in the intramolecular absorption spectra. The significant value of ρ even in the N-phase (*virtual transfer* of electronic charge from D to A *in the ground state*), is the origin of the name “charge transfer” compound. However, there is a peak in the visible region of the absorption spectrum in the N-phase which is ascribed to an optical transition corresponding to the *real* transfer of an electron from a D^0

molecule to a neighboring A^0 molecule in the zeroth-order approximation, while the corresponding peak in the I-phase is split into two components.²⁶ This splitting will be related to the presence of *free spin* on A^- and D^+ in the I-phase which is absent in the N-phase. The N–I transition is also accompanied by *dimerization* – a structural transition such that the D–A distance a_1 and the next A–D distance a_2 along a chain are different in the I-phase, in contrast to $a_1 = a_2$ in the N-phase.^{27,28}

We present here a theory which (i) describes exactly the long-range part of the Coulomb interactions among the ions and electrons (the point-charge model) and then (ii) explores the effect of intermolecular electron transfer energy t on the N–I transition and the associated discontinuity of the degree ρ of charge transfer using the *unrestricted* Hartree–Fock (UHF) approximation.²⁹ The dimerization arises in the I-phase due to the electrostatic instability in stage (i), as will be seen.

Let us consider a chain of N pairs of D^{2+} and A^0 alternately arrayed with an equal distance a along which $2N$ electrons can move through the highest occupied D orbitals and the lowest unoccupied A orbitals. The electronic–ionic part H_e of our Hamiltonian is written as

$$\begin{aligned} H_e = & \sum_{m=1}^{2N} \varepsilon_m n_m + \sum_{m=1}^{2N} U_m n_{m\uparrow} n_{m\downarrow} \\ & + \sum_m \sum_{m' < m} V_{mm'} (2v(m) - n_m)(2v(m') - n_{m'}) \\ & - \sum_{\sigma} \sum_{m=1}^{2N} t_{m,m-1} (a_{m-1\sigma}^\dagger a_{m\sigma} + a_{m\sigma}^\dagger a_{m-1\sigma}) \end{aligned} \quad (13.4.1)$$

where $a_{m\sigma}$ and $n_{m\sigma} = a_{m\sigma}^\dagger a_{m\sigma}$ are, respectively, the annihilation and number operators of an electron with spin σ at the m th site, and $n_m = \sum_{\sigma} n_{m\sigma}$. Depending on whether m is odd or even, the ionic charge $2v(m)$ takes the value 2 or 0, the on-site inter-electron Coulomb energy U_m the value U_D or U_A , and the site energy ε_m the value $-I_+$ or $-A$, where $I_+ = I + U_D$ represents the ionization energy of D^+ . Of the possible structural changes, we consider only uniform dimerization: $a \rightarrow a \pm aq$, whereby the intersite Coulomb energy is changed to $V_{mm'} = V|(m - m') + (v(m) - v(m'))q|^{-1}$ where $V = e^2/a$. Here q is an example of an order parameter mentioned in the Overview. The corresponding change in the transfer energy $t_{m,m-1}$ will be neglected and equated to t_0 since t_0 itself is small compared to V . The dimerization also causes a change in the short-range intermolecular interactions which will be considered up to the fourth-order term

$$H_V = N M_r \omega_N^2 (aq)^2 / 2 + N c (aq)^4 / 4 \quad (13.4.2)$$

where M_r is the reduced mass of D and A molecules.

Neglecting t_0 (point-charge model) for the moment and taking the completely neutral phase $N^{(0)}$ as the origin of energy H_e , one finds that the adiabatic energy of

the completely ionized phase $I^{(0)}$ is given by

$$E_I^{(0)}/N = \Delta - V \left[\alpha_M + \sum_{v=1}^{\infty} q^{2v} f_v(0) \right] + M_t \omega_N^2 a^2 q^2 / 2 + ca^4 q^4 / 4 \quad (13.4.3)$$

where α_M is the Madelung constant and

$$f_v(k) = \sum_{\ell=-\infty}^{+\infty} |2\ell - 1|^{-2v-1} \exp[-i(2\ell - 1)ka]. \quad (13.4.4)$$

Considering terms up to the fourth-order of q , one finds that the lattice is destabilized when $\eta \equiv 2V f_1(0)/(M_t \omega_N^2 a^2) > 1$, being dimerized to a finite value q_I with a new vibrational frequency $\omega_I = \omega_N [2(\eta - 1)]^{1/2}$ provided $\zeta \equiv 4V f_2(0)/(ca^4) < 1$. The electrostatic instability index readily exceeds unity in realistic situations as will be seen later. The $N^{(0)}-I^{(0)}$ boundary which was given by $\Delta - \alpha_M V = 0$ in the above is now slightly shifted so as to favor the $I^{(0)}$ phase.

For a finite transfer energy, it is more convenient to introduce the donor and acceptor Bloch orbitals, the corresponding electron creation operators being given by

$$\begin{aligned} b_{k\sigma}^\dagger &= N^{-1/2} \sum_{\ell=1}^N a_{2\ell-1,\sigma}^\dagger \exp[i(2\ell - 1)ka] \\ c_{k\sigma}^\dagger &= N^{-1/2} \sum_{\ell=1}^N a_{2\ell,\sigma}^\dagger \exp[i2\ell ka]. \end{aligned} \quad (13.4.5)$$

In order to study the ground state and the one-electron excited states of this many-electron system, we use the unrestricted Hartree–Fock approximation considering the possible existence of sublattice magnetization in the I-phase where each site has free spin in the zeroth approximation $I^{(0)}$. For the one-electron state, we use the partial charge transfer band defined by

$$d_{k\sigma}^\dagger = u_{k\sigma} b_{k\sigma}^\dagger + v_{k\sigma} c_{k\sigma}^\dagger, \quad |u_{k\sigma}|^2 + |v_{k\sigma}|^2 - 1 = 0. \quad (13.4.6)$$

The mixing coefficients $u_{k\sigma}$ and $v_{k\sigma}$ are determined so as to minimize the expectation value of the ground-state energy: $E_e = \langle G | H_e | G \rangle$. The ground state is given by $|G\rangle = \prod_{k\sigma}^{2N} d_{k\sigma}^\dagger |0\rangle$ where the product on $(k\sigma)$ is extended over the one-electron states which are occupied according to the Pauli principle.

The extremization of E_e with respect to $u_{k\sigma}$ and $v_{k\sigma}$ with Lagrange's indeterminate coefficients $\varepsilon_{k\sigma}$ for their normalization condition leads to a two-row eigenequation for $\varepsilon_{k\sigma}$, the Hartree–Fock one-electron energy. The two eigenvalues $\varepsilon_{k\sigma}^\pm$, representing the upper- and lower-band energies respectively, are given as follows where the origin of energy is put at $-I$, the one-electron energy of D^0 :

$$\varepsilon_{k\sigma}^\pm = (1/2) \left[\Delta - (U_D - U_A) \rho_{-\sigma} \pm (B_{k\sigma}^2 + |C_{k\sigma}|^2)^{1/2} \right], \quad (13.4.7)$$

$$B_{k\sigma} = B_\sigma - 2V G_\sigma(k), \quad (13.4.8)$$

$$B_\sigma = \Delta + \left[U - 2V \left\{ \alpha_M + \sum_{v=1}^{\infty} q^{2v} f_v(0) \right\} \right] \rho - U s_\sigma, \quad (13.4.9)$$

$$\rho = \rho_\uparrow + \rho_\downarrow \quad (\rho_\sigma = \langle n_{2\ell,\sigma} \rangle), \quad s = \rho_\uparrow - \rho_\downarrow \quad (13.4.10)$$

$$G_{\sigma}(k) = N^{-1} \sum_{k'} g(k - k') |v_{k'\sigma}|^2, \quad (13.4.11)$$

$$g(k - k') = \sum_{\ell=-\infty}^{\infty} |2\ell|^{-1} \exp[-2i\ell(k - k')a], \quad (13.4.12)$$

$$C_{k\sigma} = 4t_0 \cos(ka) + 2V \sum_{\nu=0}^{\infty} q^{2\nu} F_{\nu\sigma}(k), \quad (13.4.13)$$

$$F_{\nu\sigma}(k) = (2N)^{-1} \sum_{k'} f_{\nu}(k - k') y_{k'\sigma}. \quad (13.4.14)$$

Here the summation on k' extends over all occupied electron states. As mentioned in connection with UHF, we assume that the sublattice spin s vanishes in the N-phase and does not vanish in the I-phase, which is the only constraint imposed on the self-consistent solutions. As a result, $\varepsilon_{k\sigma}$ depends on σ only in the I-phase. The eigenequations for all $(k\sigma)$ should be solved self-consistently since the eigenequation for $(k\sigma)$ depends on the occupancy of other $(k'\sigma')$ which in turn depends on their one-electron energies $\varepsilon_{k'\sigma'}$ which are yet to be determined. The resulting nonlinear integral equation for $y_{k\sigma} = 2u_{k\sigma}^* v_{k\sigma}$ is given by

$$|B_{k\sigma}| y_{k\sigma} (1 - |y_{k\sigma}|^2)^{-1/2} = C_{k\sigma}. \quad (13.4.15)$$

This integral equation is solved numerically by setting $N = 80$ for the k -summation in order to avoid the finite-size effect, while the ℓ -summation for the intersite Coulomb energies are considered up to a greater distance.

Taking TTF-CA as an example, we put $a = 0.363$ nm and $V = 3.97$ eV, $M_r = 1.85 \times 10^{-25}$ kg, the LO-mode angular frequency $\omega_N = 5.65 \times 10^{12}$ s $^{-1}$ obtained by neutron scattering³⁰ in the N-phase. Noting that $f_1 = 2.1036$, one finds $\eta = 3.44$, which satisfies the criterion: $\eta > 1$. That is, the I-phase is already subject to electrostatic instability towards dimerization within the point-charge model. The degree of dimerization in the I-phase is $q_I \sim 0.03$ according to neutron²⁸ and X-ray³¹ diffraction studies. However, for the technical reason of making it feasible to distinguish the adiabatic potentials of the N- and I-phases within the precision of the numerical calculation, we put $q_I = 0.1$. One then obtains $\zeta = 2.62 \times 10^{-2}$ (note that $f_2 = 2.0095$). It seems reasonable to put $U \equiv (U_D + U_A)/2 = 1.5 \times V$. Other energies such as Δ and t_0 will be measured in units of V . Using the point-charge model and the above-mentioned parameter values, ω_I is estimated as $\sim 1.25 \times 10^{13}$ s $^{-1}$. It is consistent with the experimental results²⁸ that ω_I is significantly larger than ω_N .

Calculated adiabatic energies for the N- and I-phases are shown in Fig. 13.8 for typical sets of (Δ, t_0) -values on both sides of the N–I boundary shown by the solid lines on the phase diagram of Fig. 13.9. As seen from Fig. 13.8, the I-phase is always dimerized while the N-phase is not for small t_0 . However, the N-phase also is dimerized for large enough t_0 (Fig. 13.8(b)) because of the significant value of ρ . The boundary for the second-order transition to the dimerized N-phase is shown

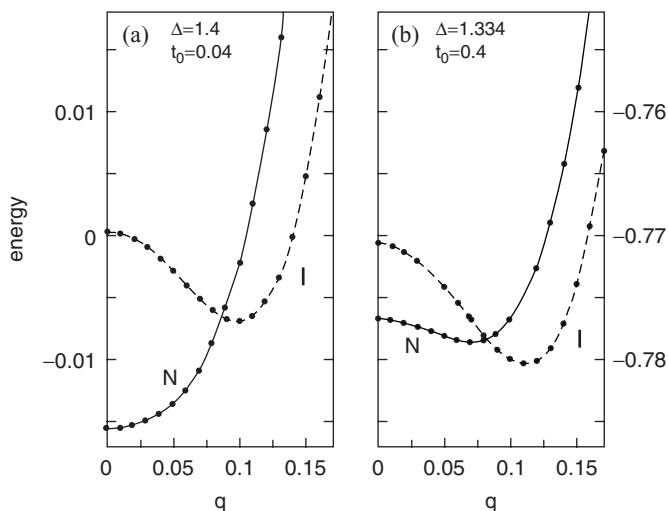


Fig. 13.8 Calculated adiabatic energies for the N-phase and the I-phase for different sets of (Δ, t_0) -values (within the metastability region of Fig. 13.9).²⁹

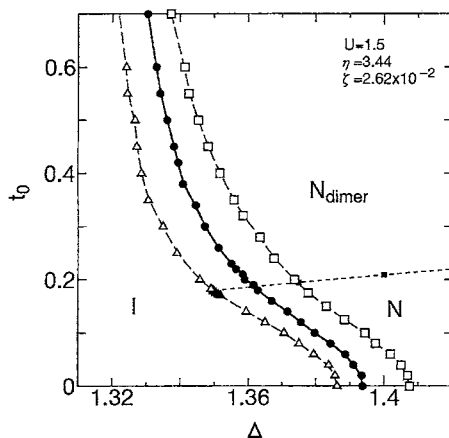


Fig. 13.9 Phase diagram for the first-order N–I transition (solid line) with metastability regions on both sides (bounded by broken lines) and for the second-order transition (dotted line) to dimerization of the N-phase.²⁹

by the dotted line in Fig. 13.9. We note in Fig. 13.8 that for any fixed value of q the lower branch of the two adiabatic curves represents the stable electronic state to be realized since the UHF is based on the variational principle. As we increase q , always tracing the lower branch, we get two minima, one before and one after the N–I crossover. Namely, we have one stable and one metastable state in the case of Fig. 13.8. This “metastability” is realized, on the phase diagram of Fig. 13.9,

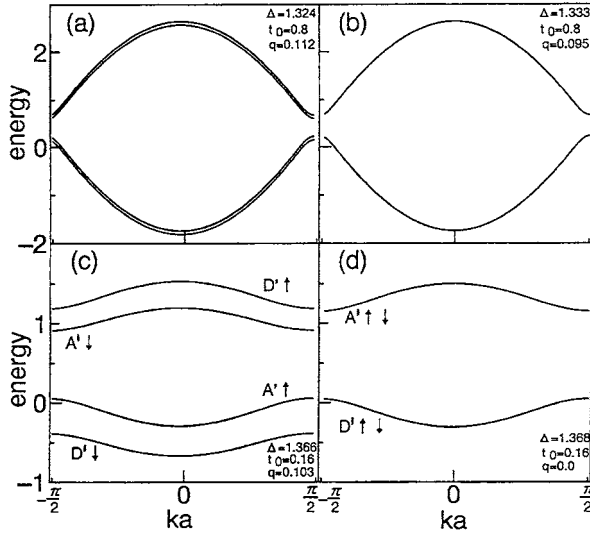


Fig. 13.10 Sub-band energies vs. k calculated with $U_D = 2$ and $U_A = 1$ for the N (right) and I-(left) phases, for small (lower) and large (upper) values of t_0 .²⁹

only within the region between the two broken lines, one with open squares and the other with open triangles. Outside this region there is only the stable phase, N or I.

Figure 13.10 shows the sub-band energies $\varepsilon_{k\sigma}^{(+)}$ against k , for the N-(right) and I-(left) phases at their adiabatic minima and for small (lower half) and large (upper half) values of t_0 . Here D'(A') denotes the mixed band with D(A)-predominance. In the I-phase, $\varepsilon_{k\sigma}^{(+)}$ depends on σ in accordance with our choice of sign: $s > 0$ on the A sublattice. In either phase, two lower bands are occupied and two upper bands unoccupied, each being spin degenerate in the N-phase. The occupancy of D'_\downarrow and A'_\uparrow bands in the I-phase corresponds in r -space to the antiferromagnetic ordering between the D and A sublattices. As is well known, the lowest magnetic excitations from such a ground state are spin waves with energy vanishing with wave number k . In contrast, the energy gap for optical excitation conserving spin and wave number is significant in both phases. The bifurcation of the optical gap in the I-phase noted in Fig. 13.10(a) and (c) will be related to observations later on. The energy width of each sub-band increases with t_0 , as expected. However, each of the two lower and two upper bands in the I-phase tends to be degenerate ((c) to (a) in the figure), approaching the spin-independent bands in the N-phase (b) asymptotically.

The three quantities, ρ , s and q_{\min} , are subject to discontinuous changes as Δ decreases across the N–I transition line. Comparing the observed value $\rho_N \sim 0.3$ in TTF-CA just above the transition temperature with the calculated ρ_{N-t_0} curve, one

finds $t_0 \sim 0.15 (\sim 0.6 \text{ eV})$, which is larger than was estimated by other groups,³² but still falls below the dotted line in Fig. 13.9.

The results of the above calculation are consistent with experimental results on TTF-CA with a discontinuous increase in ρ at the N–I transition and the presence of dimerization and electron spin resonance (due to the spin solitons³³ associated with the vanishing magnetic gap) in the I-phase only. Also, the existence of *two* optical gaps with different energies in the I-phase ($A'_\uparrow \rightarrow D'_\uparrow$ and $D'_\downarrow \rightarrow A'_\downarrow$ in Fig. 13.10(c)) in contrast to a *single* optical gap in the N-phase ($D'_\sigma \rightarrow A'_\sigma$ in Fig. 13.10(d)) is consistent with the observation of two exciton absorption peaks in the I-phase in contrast to a single peak in the N-phase.²⁶ As regards the presence or absence of dimerization in the two phases, it is also of interest to compare the phase diagram Fig. 13.9 with observations on TMB-TCNQ.³⁴

While the application of the mean field theory to the phase transition in one-dimensional systems is less controversial for long-range interaction, it seems to predict that the N–I transition line persists up to too large a value of t_0 . Further study with intrachain fluctuations and interchain coupling is desirable.

13.5 Photo-induced phase transition in TTF-CA

It was first shown by Koshihara *et al.*³⁵ that excitation of the TTF-CA crystal below the transition temperature T_c with light in the visible region induces the transition from the ionic to the neutral phase, over a lattice region as large as *hundreds of unit cells per one photon*. By careful examination of the change in the reflectance spectrum, they concluded that the transition is not associated with any temperature increase during photoexcitation but is induced by purely electronic stimulation. In analogy with the fact that the site substitutional impurities with lower electronic affinity than CA can induce a semi-macroscopic neutral domain in the I-phase,³⁶ they considered this photo-induced phase transition (PIPT) from the I- to N-phase to be due to *photo-injection* of neutral DA pairs into the I-phase. Correspondingly, the photo-induced N- to I-phase transition of the crystal above T_c was also observed. In either case, the photo-induced transformation due to a pulsed light ($\sim 80 \text{ fs}$) takes a time $\tau \sim \text{hundreds ps}$ as seen in Fig. 13.11, and the metastable phase remains for an order of a millisecond. The dependence of the converted fraction on the intensity of the excitation light shows a typical *threshold* behavior, indicating the *cooperative* nature of the transformation process. See Refs. [37] and [38] for the reviews of experimental and theoretical studies on PIPT and related phenomena in different materials. According to a recent study,³⁹ the threshold intensity in TTF-CA depends on the photon energy of the excitation light, being negligibly small for higher energy as shown in Fig. 13.12.

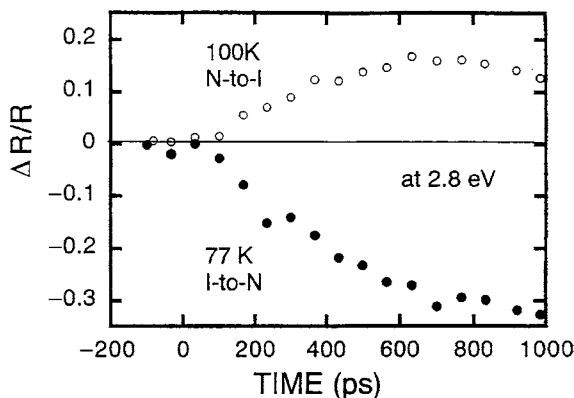


Fig. 13.11 Time profiles of photorefectance observed at 2.8 eV induced by irradiation with a 1.55 eV light pulse with 80 fs pulse width. The photon flux for excitation was $1 \times 10^4 \text{ cm}^{-2}$. Due to Koshihara *et al.*³⁷

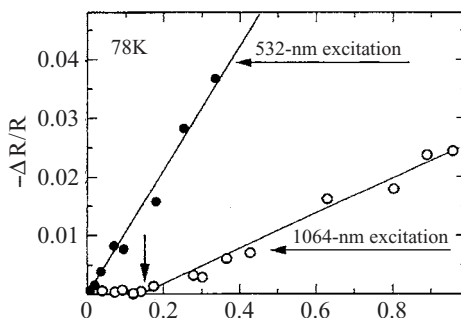


Fig. 13.12 Magnitude of the reflectance changes at 3.0 eV induced by 1064 nm and 532 nm laser pulses, as a function of the excitation intensity of the laser pulses scaled in terms of the fluence (total number of photons which have passed a unit area). Due to Suzuki *et al.*³⁹

The motion of the neutral–ionic domain wall (NIDW),⁴⁰ the role of dipole–dipole interactions between the excited DA pairs,⁴¹ the model of excitonic strings for the cooperative aspect of the many-electron excited states⁴² and adiabatic pictures for the proliferation of charge transfer excitons leading to the formation of semi-macroscopic metastable domains⁴³ have been studied in order to depict various aspects of PIPT in TTF-CA. As a complementary approach, we present here a tentative *semi-phenomenological adiabatic* description of PIPT based on the theory described in the preceding section.

Typical adiabatic potentials $E_N(q)$ and $E_I(q)$ for the neutral and ionic states, respectively, such as were presented in Fig. 13.8 for typical values of parameters, are now shown schematically in Fig. 13.13 against the dimerization coordinate q as the abscissa, for the realistic value of t_0 (~ 0.1) and in the low-temperature

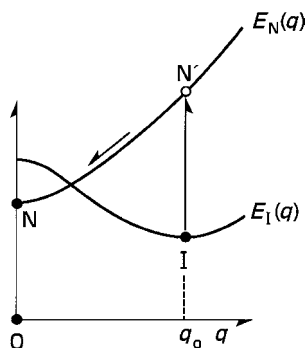


Fig. 13.13 Adiabatic potentials $E_I(q)$ and $E_N(q)$ against dimerization coordinate q for a realistic value of t_0 in the low-temperature phase (I).

region in which the ionic phase with dimerization q_0 has the lowest energy. Let us normalize the ordinate as the energy *per unit cell*. In accordance with the argument in Section 4.4, the minima of the adiabatic energies, $E_N(0)$ and $E_I(q_0)$, are to be considered to represent *free energies* (as is obvious from the fact that the temperature-dependent lattice constant a and the dimerization coordinate q are treated as phenomenological variables governing $E_N(q)$, $E_I(q)$ and hence the N–I transition temperature in the theory of Section 13.4) per unit cell of the two phases, whose difference is on the order of the thermal energy and approaches zero as T tends to T_c . This is the obvious reason why so many unit cells are transformed into the metastable phase per one photon ($h\nu \sim \text{eV}$). According to the Franck–Condon principle, the electronic transition between the ionic and neutral states in the configuration coordinate space should take place vertically, keeping the coordinate q (and the parameter a) fixed (the upward arrow in Fig. 13.13). Even at q_0 , the energy difference $E_N(q_0) - E_I(q_0)$ is much smaller than the photon energy $h\nu$ as mentioned above. Within our *phenomenological* adiabatic model with *two levels*, I and N, the *real* (energy-conserving) absorption of a photon can take place only over a *semi-macroscopic* region consisting of $m_1 (\gg 1)$ unit cells such that

$$m_1[E_N(q_0) - E_I(q_0)] + E_{\text{DW}} = h\nu, \quad (13.5.1)$$

where we have added the energy E_{DW} needed to form a pair of NIDWs. This corresponds microscopically to simultaneous excitation of successive m_1 pairs of D^+A^- into D^0A^0 states (or m_1 pairs of (inverse) charge transfer excitons).

Such a multi-electron transition by a single photon would be extremely improbable due to the weak electron–photon interaction *but for a strong enough* electron–electron interaction which would promote a cooperative excitation. There exists in fact a certain ground to argue for the latter situation. The excitation energy to bring a single D^+A^- pair into the D^0A^0 state in the matrix of ionized pairs,

given by $(2\alpha_M - 1)e^2/a - \Delta(> 0)$ (under the neglect of the dimerization effect), is much larger than $E_N(q_0) - E_I(q_0)$ which is on the order of the thermal energy as mentioned above. It is the long-range electrostatic interactions (< 0) between dipoles caused by back-transferred charges (from A^- to D^+) on m_1 pairs that make the total energy, needed to transform the electronic states of the region of m_1 unit cells from I to N, as small as the l.h.s. of eq. (13.5.1). Namely, the interaction energies are so large and long ranged that they as a whole almost cancel the sum of the individual excitation energies. Therefore, the coupling constant, defined by the ratio of the interaction energies to the individual energies, is of the order of unity, making multi-electron excitations as probable as individual excitations. This seems to be the reason why the exciton string model was introduced, and why one can expect the cooperative excitation of many electrons to be possible thereby.

On the other hand, the dipole-dipole interactions are already included in $[E_N(q_0) - E_I(q_0)]$ calculated in Ref. [29] where the long-range Coulomb interaction was exactly considered. So we simply *assume* that for a given photon energy $h\nu$ the entire region of m_1 unit cells (with m_1 being given by (13.5.1)) *can make a cooperative optical transition* from I to N with a *reasonable finite* rate, leaving unspecified how one can calculate the rate of this multi-electron excitation process by a single photon (in terms of say, a higher-order perturbation expansion in the electron-electron interaction).

If we once assume a finite matrix element for this optical transition shown by the upward arrow in Fig. 13.13, we can resort to the conventional C. C. model in describing the lattice relaxation after the optical transition. The relaxation takes place along the adiabatic potential $E_N(q)$ from $q = q_0$ to $q = 0$, as shown by the down-hill arrow. Because of weak interchain interactions, the energy released thereby can only be transformed to extend the N-phase region from m_1 to m'_1 unit cells within the same chain, as given by

$$m'_1[E_N(0) - E_I(q_0)] + E_{DW'} = m_1[E_N(q_0) - E_I(q_0)] + E_{DW}, \quad (13.5.2)$$

where $E_{DW'}$ in the lattice-relaxed state is smaller than E_{DW} . The dipole-dipole interaction energies in the relaxed state, as was discussed in Refs. [41, 44], are included in $E_N(0)$, while those in the unrelaxed state are considered in the exciton strings⁴² and included in $E_N(q_0)$.

For P photons absorbed per unit length consisting of N_0 unit cells, the fractional volume of the metastable N-phase is given by

$$f_N^1 = m'_1 P / N_0 = (P / N_0)[h\nu - E_{DW'}] / [E_N(0) - E_I(q_0)]. \quad (13.5.3)$$

This result holds only for $h\nu > E_{DW}$ when one-photon absorption is possible.

If $(E_{\text{DW}}/2) < h\nu < E_{\text{DW}}$, we have to consider the event that two photons are absorbed at unit cells within the mutual distance m_2 of the size of the transformed region given by

$$m_2[E_N(q_0) - E_I(q_0)] + E_{\text{DW}} = 2h\nu. \quad (13.5.4)$$

The probability for this event is given by $P(2m_2/N_0)$ provided the latter is much smaller than unity. Since the number of such pairs is given by $(1/2)P^2(2m_2/N_0)$, the fractional number of the metastable N -phase is given by

$$f_N^2 = (m_2'/N_0)(P^2 m_2/N_0) = (P/N_0)^2 (2h\nu - E_{\text{DW}})[E_N(q_0) - E_I(q_0)]^{-1} \\ \times (2h\nu - E_{\text{DW}})[E_N(0) - E_I(q_0)]^{-1} \quad (13.5.5)$$

where m_2' is given by (13.5.2) with m_1 and m_1' being replaced by m_2 and m_2' . Equations (13.5.3) and (13.5.5) seem to qualitatively reproduce the two curves in Fig. 13.12 for high- and low-energy photons, except that the threshold behavior for the low-energy photon in the latter is represented by the quadratic curve in the former, a discrepancy which could be overcome only by considering the many-body aspect of the optical transition more appropriately. The energy E_{DW} is estimated to be between 1.17 eV and 2.33 eV, which is not unreasonable. The assumed existence of a potential barrier due to E_{DW} in the case of excitation with a low-energy photon is consistent with the microscopic study of the adiabatic potential.⁴³

The time τ needed for lattice relaxation is governed by the sound velocity in the chain with which NIDW proceeds. In order for the NIDW to proceed a distance 300 times the lattice constant $2a(= 0.73 \text{ nm})$, as observed, with the sound velocity estimated as $0.5 \times 10^3 \text{ m s}^{-1}$, it takes $\tau \sim 0.4 \text{ ns}$, which is consistent with observation.

13.6 T - U - S model for N -electron N -site systems

We shall now be concerned with one of the most ubiquitous systems among molecules and solids, those consisting of N electrons and N equivalent sites, which we have considered relatively little in the preceding sections in spite of their importance. One can find such systems with a great variety in size, dimensionality and property, including the simplest, the hydrogen molecule ($N = 2$), up to monovalent metals of macroscopic size ($N = \infty^3$), as will be mentioned later. Our main concern is how the electronic behaviors of such systems are *governed by the competition of the three independent elements of interaction* which we have already introduced in different contexts, (i) T : the energy of electron transfer between nearest neighboring (n.n.) sites (T , denoted thus far by the lower case t , is assumed to be positive), (ii) U : the Coulomb repulsive energy of two electrons on the same site, (iii) S : the

on-site electron–phonon interaction energy, defined here only as *twice* the lattice relaxation energy $E_R(= \hbar\omega)$ introduced in preceding sections.

Our model Hamiltonian is given in terms of these parameters as

$$\begin{aligned} H &= H_e + H_{ep} + H_p \\ &= -T \sum_{\ell, \ell'} \sum_{(n.n.)} \sum_{\sigma} a_{\ell\sigma}^{\dagger} a_{\ell'\sigma} + U \sum_{\ell} n_{\ell\uparrow} n_{\ell\downarrow} \\ &\quad - S^{1/2} \sum_{\ell} \sum_{\sigma} Q_{\ell} n_{\ell\sigma} + \sum_{\ell} (Q_{\ell}^2 + \omega^2 P_{\ell}^2)/2, \end{aligned} \quad (13.6.1)$$

where $a_{\ell\sigma}$ denotes the annihilation operator for an electron with (z -component of) spin σ (\uparrow or \downarrow) at the ℓ th site, and $n_{\ell\sigma} = a_{\ell\sigma}^{\dagger} a_{\ell\sigma}$ is the corresponding number operator. The components of the spin s_{ℓ} at the ℓ th site are given by

$$\begin{aligned} s_{\ell x} &= (1/2)(a_{\ell\uparrow}^{\dagger} a_{\ell\downarrow} + a_{\ell\downarrow}^{\dagger} a_{\ell\uparrow}), \\ s_{\ell y} &= (-i/2)(a_{\ell\uparrow}^{\dagger} a_{\ell\downarrow} - a_{\ell\downarrow}^{\dagger} a_{\ell\uparrow}), \\ s_{\ell z} &= (1/2)(n_{\ell\uparrow} - n_{\ell\downarrow}). \end{aligned} \quad (13.6.2)$$

The r.h.s. of (13.6.1) without the second term is the Hamiltonian for the electron–phonon system which results in self-trapping of an electron for large S against T in a one-electron system as was described in Section 9.4 in a somewhat different form. The first line of (13.6.1) is the so-called Hubbard Hamiltonian which describes⁴⁵ the transition from metal to insulator in a solid with increasing U against T and the gradual transition from Hund–Mulliken (molecular orbital) to Heitler–London pictures in the hydrogen molecule, as shown on the triangular coordinate diagram in Fig. 13.14. (A point on this diagram represents the parameter values T , U and S which are proportional to their distances to US -, ST - and TU -sides.) In

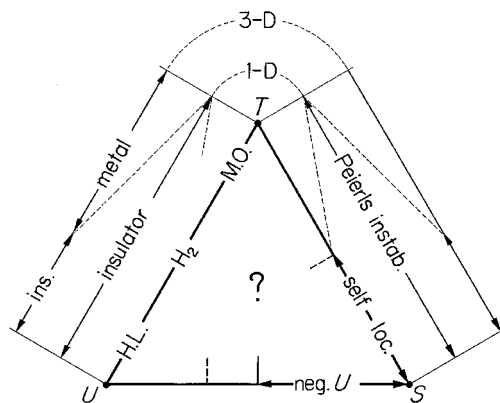


Fig. 13.14 T – U – S triangular coordinate with typical 0-, 1- and 3-dimensional problems of competition of two out of T , U and S .⁴⁶ Here T represents the intersite transfer energy, U the intrasite Coulomb repulsive energy and S twice the lattice relaxation energy.

a one-dimensional lattice, the Peierls transition to an insulating phase with spin- or charge-density wave takes place with *infinitesimally* small U or S against T . Along the bottom of the triangle, we have the negative U effect in the system of two electrons on two distant sites ($T = 0$) as we increase S against U . The superconducting state follows from (13.6.1) when the phonon-mediated attraction between electrons in the k -space overcomes the Coulomb repulsion. Although a finite phonon energy $\hbar\omega$ plays an essential role in the last-mentioned state, we shall confine ourselves to the adiabatic limit neglecting $\hbar\omega$ against the other three energy parameters, T , U and S , the competition between which is our concern here.

Let us introduce *an orthogonal* transformation in the N -dimensional space, $\{Q_\ell\} \rightarrow \{q_j\}$, $\{P_\ell\} \rightarrow \{p_j\}$ and $\{n_{\ell\sigma}\} \rightarrow \{v_{j\sigma}\}$ defined by

$$Q_\ell = \sum_j c_{\ell j} q_j, \quad P_\ell = \sum_j c_{\ell j} p_j, \quad n_{\ell\sigma} = \sum_j c_{\ell j} v_{j\sigma} \quad (13.6.3)$$

which keeps the following scalar products invariant:

$$\sum_\ell Q_\ell^2 = \sum_j q_j^2, \quad \sum_\ell Q_\ell n_{\ell\sigma} = \sum_j q_j v_{j\sigma}, \quad \text{etc.} \quad (13.6.4)$$

For the validity of (13.6.4) the coefficients must satisfy the orthonormality relations:

$$\sum_\ell c_{\ell j} c_{\ell j'} = \delta_{jj'}. \quad (13.6.5)$$

The reciprocal transformation of (13.6.3) is given, with the use of (13.6.5), by

$$q_j = \sum_\ell c_{\ell j} Q_\ell, \quad \text{etc.} \quad (13.6.6)$$

Let us now specify the first of the new coordinates as

$$q_1 \equiv N^{-1/2} \sum_\ell Q_\ell, \quad \text{etc., namely } c_{\ell 1} = N^{-1/2}. \quad (13.6.7)$$

The up-down sum of the corresponding number operators $v_{1\sigma}$ turns out to be a c-number:

$$v_1 \equiv \sum_\sigma v_{1\sigma} \equiv N^{-1/2} \sum_\ell \sum_\sigma n_{\ell\sigma} = N^{1/2}. \quad (13.6.8)$$

Therefore, in the new expression

$$H_{\text{ep}} = -S^{1/2} \sum_j \sum_\sigma q_j v_{j\sigma}, \quad (13.6.9)$$

the $j = 1$ term does not depend on the electronic state. Combining this term with the $j = 1$ term of the new expressions of H_p , one obtains a harmonic oscillator

$$H_1 = [\{q_1 - (NS)^{1/2}\}^2 + \omega^2 p_1^2]/2 - NS/2 \quad (13.6.10)$$

which is independent of the electronic state. One can thus decouple (13.6.10) from (13.6.1) to obtain an effective Hamiltonian:

$$H_{\text{eff}} \equiv H - H_1 = H_e + H_{\text{ep}}' + H_{\text{p}}', \quad (13.6.11)$$

$$H_{\text{ep}}' \equiv -S^{1/2} \sum_{j(\neq 1)}' \sum_{\sigma} q_j v_{j\sigma}, \quad (13.6.12)$$

$$H_{\text{p}}' \equiv \sum_{j(\neq 1)}' [q_j^2 + \omega^2 p_j^2]/2. \quad (13.6.13)$$

Since the total spin defined by $s = \sum_{\ell} s_{\ell}$ commutes with the Hamiltonian (13.6.1), one can specify s and s_z in solving the eigenvalues of H . Let us consider for reference the state with highest possible spin and its z -component: $s = s_z = N/2$, in which $n_{\ell\uparrow} = 1$ and $n_{\ell\downarrow} = 0$ for all ℓ . Due to the spin constraint, the orbital motion is completely quenched. In fact, $\langle H_e \rangle = 0$ since $\langle a_{\ell\sigma}^{\dagger} a_{\ell\sigma} \rangle = 0$ and $\langle n_{\ell\downarrow} \rangle = 0$, and $\langle H_{\text{ep}}' \rangle = 0$ since $\langle v_{j\sigma} \rangle = \sum_{\ell} c_{\ell j} \langle n_{\ell\sigma} \rangle = \delta_{\sigma\uparrow} \sum_{\ell} c_{\ell j} = N^{1/2} \delta_{j1} \delta_{\sigma\uparrow}$ where we have multiplied the factor $1 = N^{1/2} c_{\ell 1}$ (see (13.6.7)) to each summand on the l.h.s. of the last equation to obtain the r.h.s. Therefore, the origin of electronic energy in H_{eff} is at the state of the highest spin, $S = N/2$, where the orbital motion is completely quenched.

In the following, we will begin with the simplest system and extend it to larger systems with a variety of structures.

13.6.1 The 2-electron 2-site system⁴⁶

The only mode q_2 which interacts with the electrons and the corresponding electron number v_2 are given, with the use of orthonormality (13.6.5) to the totally symmetric mode q_1 of (13.6.7), by

$$q_2 \equiv 2^{-1/2}(Q_1 - Q_2), \quad v_2 \equiv 2^{-1/2} \sum_{\sigma} (n_{1\sigma} - n_{2\sigma}), \quad (13.6.14)$$

and the effective Hamiltonian by

$$\begin{aligned} H_{\text{eff}} &= H_e - S^{1/2} v_2 q_2 + (q_2^2 + \omega^2 p_2^2)/2 \\ &\equiv H(q_2) + (q_2^2 + \omega^2 p_2^2)/2. \end{aligned} \quad (13.6.15)$$

Namely, the charge transfer v_2 interacts only with the antisymmetric mode q_2 , as it should. Hereafter, we will confine ourselves to the adiabatic limit: $\omega \rightarrow 0$.

As for the spin, there is one triplet (t) $S = 1$ and three singlets (s) $S = 0$. The triplet state, with each site being occupied by one electron, has $H_e = 0$ as mentioned before in the general case. The singlet consists of homopolar (h) and ionized states (i):

$$\begin{aligned} |\text{h}\rangle &= 2^{-1/2}(a_{1\uparrow}^{\dagger} a_{2\downarrow}^{\dagger} + a_{2\uparrow}^{\dagger} a_{1\downarrow}^{\dagger})|0\rangle, \\ |\text{i}\rangle &= a_{i\uparrow}^{\dagger} a_{i\downarrow}^{\dagger}|0\rangle, \quad (\text{i} = 1, 2) \end{aligned} \quad (13.6.16)$$

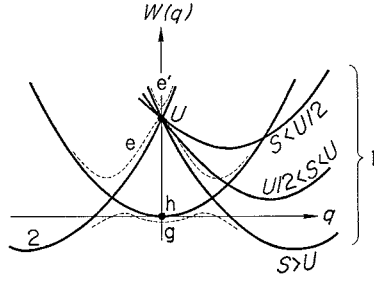


Fig. 13.15 Adiabatic potentials for vanishing (solid lines) and small- T (dashed lines) values of the 2-electron 2-site system.⁴⁶

with the corresponding energy matrix:

$$H^{(s)}(q_2) : \begin{array}{c} \begin{array}{ccc} |h\rangle & |1\rangle & |2\rangle \\ \hline \begin{bmatrix} 0 & -2^{1/2}T & -2^{1/2}T \\ -2^{1/2}T & U - S^{1/2}q_2 & 0 \\ -2^{1/2}T & 0 & U + S^{1/2}q_2 \end{bmatrix} \end{array} \end{array} \quad (13.6.17)$$

The adiabatic energy (13.6.15) for the three eigenstates of (13.6.17) with vanishing and nonvanishing T are shown in Fig. 13.15 by solid and dashed lines, respectively. With very small T , subsidiary minima (metastable configurations) of the lowest branch, which originate from the ionized states, appear for intermediate electron-phonon coupling: $U/2 < S < U$ and become more stable for strong coupling: $S > U$. With finite T and U but vanishing S or q_2 , the system is nothing other than a hydrogen molecule: (13.6.5) has eigenvalues $E_{g,e'} = [U \mp (U^2 + 16T^2)^{1/2}]/2$ and $E_e = U$, of which the ground state (g) represents the Heitler-London and the Hund-Mulliken states in the limits of $4T \ll U$ and $4T \gg U$, respectively. With vanishing U but finite T and $S^{1/2}q_2$, the corresponding eigenenergies are given by $E_{g,e'} = \mp(4T^2 + 2Sq_2^2)^{1/2}$ and $E_e(q_2) = 0$. The minimum point of the adiabatic potential $W_e(q_2) = E_g(q_2) + q_2^2/2$ bifurcates from $q_2 = 0$ to $q_2 = \pm[2(S^2 - T^2)/S]^{1/2}$ as S exceeds T , namely, charge transfer instability sets in at $S = T$ as a second-order transition.

Figure 13.16 is, so to speak, a phase diagram indicating, for the whole region of the (T, U, S) triangle, the stable (solid hatching) and metastable (broken hatching) configurations of the lowest electronic state, where the horizontal and vertical hatchings represent the symmetric ($q_2 = \langle v_2 \rangle = 0$) and the asymmetric ($q_2 = S^{-1/2}\langle v_2 \rangle \neq 0$) states, respectively. The adiabatic potential versus q for the ground and the lowest excited (e) states are also shown schematically for each region. Symmetry breaking in the ground state takes place as a first-order transition across the solid line

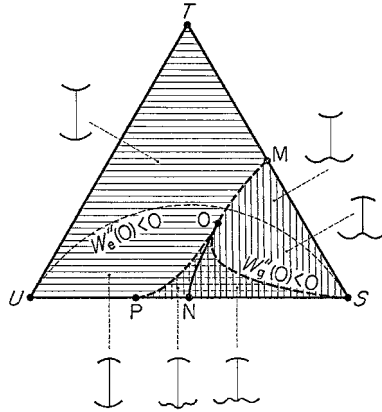


Fig. 13.16 Phase diagram for the 2-electron 2-site system. Horizontal and vertical hatchings represent symmetric and non-symmetric (charge transferred) states, respectively; solid and broken hatchings represent stable and metastable states, respectively. The adiabatic potentials for the states g and e are shown schematically for each region.⁴⁶

NO and as a second-order transition across the broken line OM. The first excited state $|e\rangle$ is also subject to second-order symmetry breaking as T decreases across the short-dashed circular arc; even a small S can cause this for $T \ll U$ because of the pseudo Jahn-Teller effect between the states e and e' which have an energy difference of $O(T^2/U)$.

In Fig. 13.15, one can see how the relaxed excited state turns out to be the real ground state in the parameter region with vertical hatching in Fig. 13.16. This is the simplest example of the situation described in Fig. 13.1, which represents the electronic-structural phase transition (of the first order) in macroscopic systems ($N \rightarrow \infty$) described later.

13.6.2 Ring systems with N -electrons N -sites

It is possible to solve exactly ring systems with not too great N . Instead of describing the individual systems, however, we note the following general features.

Let us start with the case of $U = S = 0$. Then we have one-electron Bloch states

$$N^{-1/2} \sum_{\ell=1}^N \exp(ik\ell a) a_{\ell\sigma}^\dagger |0\rangle$$

with the pseudo energy band $E(k) = -2T \cos(ka)$, where the wave number k can take values given by

$$Nak/2\pi = -(N/2) + 1, -(N/2) + 2, \dots, 0, \dots, +(N/2)$$

due to the cyclic boundary condition: the N th site is a nearest neighbor of the first site in such ring systems (a is the nearest neighbor distance). Because of the

spin degeneracy and of the orbital degeneracy $E(k) = E(-k)$ except for $k = 0$, the highest occupied level (the quasi-continuous Bloch band is up to about half filled, $E_F \sim 0$ and $k_F \sim \pi/2a = k_0/2$ when $N \gg 1$) is saturated when $N = 4M - 2$ but only doubly occupied when $N = 4M$ (M is a positive integer). In the latter case, we have orbital degeneracy of the electronic states and, hence, instability against the alternating mode $q_a \equiv N^{-1/2}(Q_1 - Q_2 + \cdots + Q_{N-1} - Q_N)$ (corresponding to the wave vector $k_F = \pi/2a$) even with an infinitesimal S , due to the Jahn-Teller (J-T) theorem. Thus the Peierls instability exists already for a finite length chain of $N = 4M$. If already U has a small but finite value, the orbital degeneracy is removed since the two uppermost electrons will occupy $\cos(kna)$ and $\sin(kna)$ orbitals ($k = k_F$) one by one to avoid the Coulomb repulsion. We then have the pseudo J-T effect, and a finite S , enough to cope with Coulomb splitting, is needed for destabilization to take place. For this reason, the boundary between the distorted and undistorted regions on the $B(= 2T)$ - U - S triangle reaches the B vertex, as shown in the lower right corner of Fig. 13.17 which is the exact result for the $N = 4$ system.⁴⁷

In contrast, the $N = 4M - 2$ systems have neither a J-T nor a pseudo J-T effect as long as M is finite. The lattice instability occurs only when S is large enough. This is the reason why the distorted region does not reach the B vertex as shown in the lower left corner of Fig. 13.17 which is a reproduction of Fig. 13.16. The

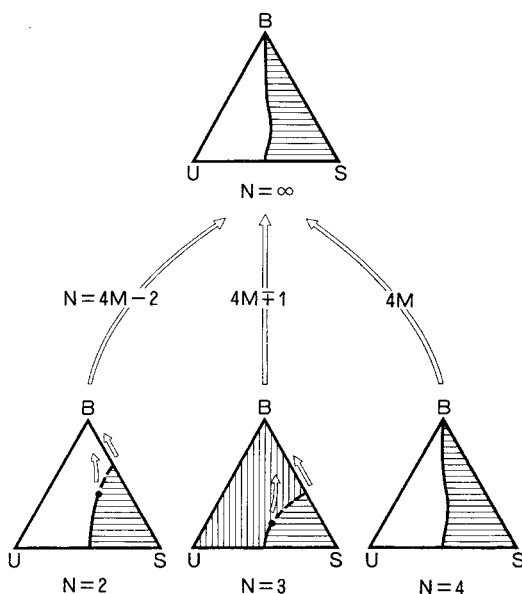


Fig. 13.17 Phase diagrams of ring-shaped N -electron N -site systems, behaving differently for three series of N .⁴⁸ Here $B = 2T$.

existence of a finite gap for electronic excitation makes these systems relatively stable in comparison with $N = 4M$ and $N = \text{odd}$ (described below) systems. This stability is known as Hückel's $4n + 2$ ($n \leftrightarrow \text{our } M$) rule in molecular science.⁴⁹ Note that we are here concerned with one π -electron per carbon atom (in addition to σ -electrons participating in the covalent bonds) in chemically saturated molecules such as aromatic molecules. As M increases, the gap decreases as $O(B/N)$ and the pseudo J-T effect due to neighboring $|k|$ s become important, making $N = 4M - 2$ systems indistinguishable from $N = 4$ systems, as they should be.

From the foregoing argument on the “closed shell” structure of $N = 4M - 2$ systems, we find that the odd- N systems ($N = 4M - 2 \pm 1$) behave similarly to each other, as one-electron or one-hole systems. The unpaired electron (or hole) has an orbital degeneracy (of sine and cosine types), which is no longer removed by introducing U since there is no opponent for mutual repulsion. This means that the region around the B vertex is subject to the J-T distortion. This is borne out by an exact calculation for the $N = 3$ ring system due to Schreiber,⁵⁰ as shown by the phase diagram in the lower center of Fig. 13.17. In conformity with the orthonormality condition (13.6.5), one can define, in addition to the totally symmetric mode q_1 (denoted by Q in Ref. [50]),

$$q_2 \equiv 6^{-1/2}(2Q_1 - Q_2 - Q_3) \quad \text{and} \quad q_3 \equiv 3^{-1/2}(Q_2 - Q_3) \quad (13.6.18)$$

(corresponding to q in Ref. [46]). The lattice instability always takes place except along the $U-B$ edge and the left half of the $U-S$ edge. A bisectric symmetry remains in the distorted 3-site ring in the region of the diagram with vertical hatching, while no symmetry remains in the region with horizontal hatching. As one increases S across the broken line (the second-order transition), each of the three equivalent J-T patterns of distortion bifurcates into two patterns of lower symmetry, tending towards the states with one and two electrons self-trapped on two sites, as shown by the calculated adiabatic potential of the ground state in the (q_2, q_3) -plane. The $N = 3$ system, like the trigonal lattice, is very interesting also in having “frustrated” spin states, which have also been studied in detail through the spin correlations: $\langle s_1 \cdot s_2 \rangle$, etc. Here, “frustration” refers to the well-known, topologically intriguing effect appearing in the system of *antiferromagnetically* coupled spins arrayed for instance on a *triangular* lattice (similar to the present system), such that any nearest pair of spins cannot take antiparallel directions in the ground state, that being against minimization of the pair energy.

Obviously, the phase diagram along the $U-S$ edge with negative U region on its right half is common to any N or even to any dimensionality since there is no interatomic transfer T .

13.6.3 One-dimensional Hubbard–Peierls system

According to Peierls,⁵¹ degenerate electrons in a one-dimensional lattice with Fermi wave number k_F are unstable against the formation of a charge-density wave (CDW) or a spin-density wave (SDW) with wavelength $\lambda = 2\pi/k_F$ even for small electron–phonon interaction or interelectron Coulomb interaction. The stabilized state is an insulator with a finite energy gap above the Fermi energy, implying that a one-dimensional metal cannot exist (Peierls theorem). Since k_F of the N -electron N -site system is given by $(L/2\pi)2k_F = N$, λ is given by $2(L/N) = 2a$, namely, the induced superlattice has a periodicity twice as large as the original one.

In conformity with this theorem, the foregoing argument on the N -electron N -site rings can be extended to an infinitely long chain ($N \rightarrow \infty$) on the assumption that the only symmetry-breaking mode of distortion or spin fluctuation realized in the lowest electronic state is the alternating mode

$$q_a = N^{-1/2} \sum_{\ell=1}^{N/2} (Q_{2\ell-1} - Q_{2\ell}), \text{ namely } Q_\ell = (2N)^{-1/2} (-1)^{\ell-1} q_a. \quad (13.6.19)$$

(see q_a introduced in Subsection 13.6.2 and also compare with eq. (13.6.14)). Here $N^{-1/2}q_a$ is another example of the order parameter mentioned in the Overview. The molecular field approximation,⁵² with further consideration of fluctuation,⁵³ has been taken for the electrons. The former exactly bisects the T - U - S triangle, with the first-order transition line, into SDW and CDW regions, while the latter has the effect of shifting the transition line slightly towards the S -side in the intermediate region, similarly to the case of $N = 4$.

In connection with this, we refer to an interesting series of spectroscopic studies on quasi-one-dimensional compounds (belonging to the Wolfram’s red salt) which consist of chains with an alternating array of Pt-ions (represented by Pt^{n+}) and halide ions (X^-) in their cores. Instead of a hypothetical equidistant array of Pt^{3+} and X^- , the real structure is such that the metal ions are alternately Pt^{2+} and Pt^{4+} , with neighboring pairs of X^- s being closer to the latter than to the former, as shown schematically in Fig. 13.18. This is nothing other than the CDW structure (in its strong limit in the sense that almost one electron is transferred from Pt^{3+} to a neighboring Pt^{3+}) formed from the above-mentioned hypothetical structure (see Fig. 13.1) by the Peierls transition with S overcoming U (corresponding to the hatched region in the top phase diagram of Fig. 13.17.) In the opposite case of $U > S$, the system would have SDW structure. The interaction mode Q_ℓ which appears in the Hamiltonian (13.6.1) represents the stretching mode of a pair of X^- ions around the ℓ th Pt ion (even-mode around the latter). The real structure is such that the alternating mode q_a of (13.6.19) has been frozen in. The resonant Raman spectrum consisting of twelve almost equidistant lines has been observed by

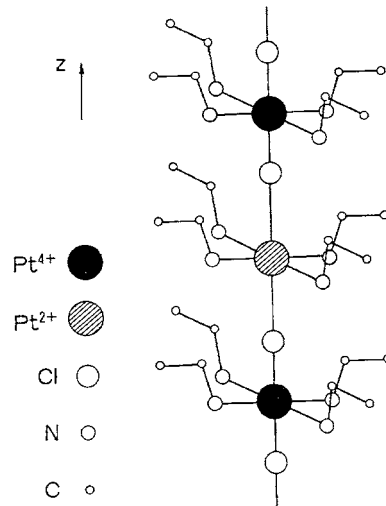


Fig. 13.18 Structure of the chain in Wolfram's red salt crystal.⁵⁴

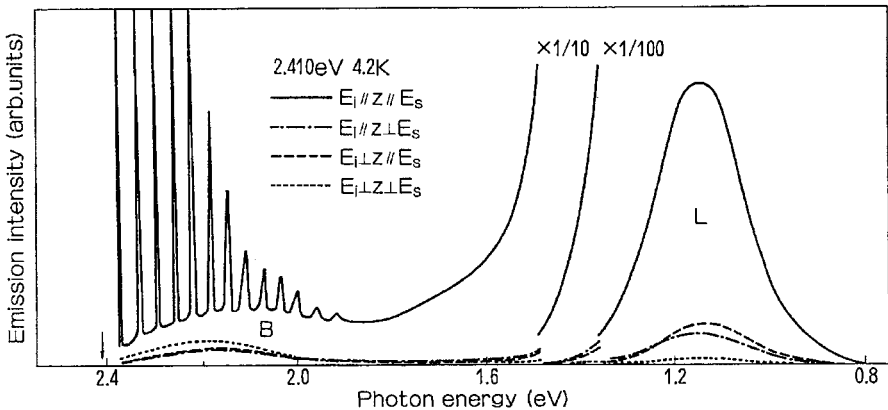


Fig. 13.19 Emission spectra (inclusive of Raman lines) of Wolfram's red salt excited at 2.410 eV where the absorption is due to the charge transfer excitation from Pt^{2+} to Pt^{4+} . Here $E_i(E_s)$ is the electric field of the incident (scattered) light and the direction of the chain is z .⁵⁴

Tanino and Kobayashi⁵⁴ for incident light resonating the exciton absorption peak and with the only polarizations of the incident and scattered lights both parallel to the chain axis (see Fig. 13.19), indicating that the stretching motions Q_ℓ of the X^- ion pairs (around the Pt ions) parallel to the chain axis are responsible for the Raman lines and the alternate mode q_a consisting of these motions is responsible for the Peierls destabilization. Moreover, the Stokes-shifted broad emission has also been observed on the low-energy side of the Raman lines, and has been ascribed to the self-trapped exciton. The exciton in this material is of the charge transfer

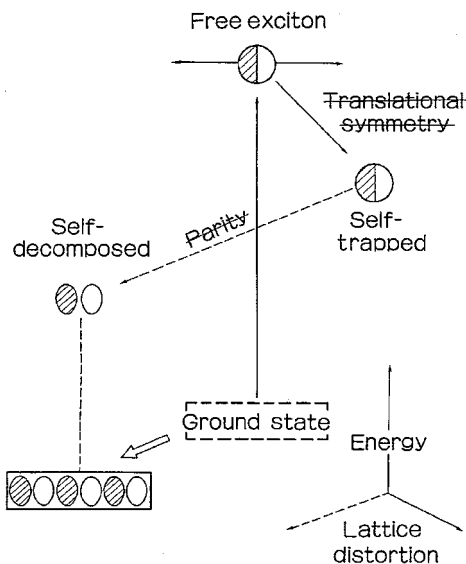


Fig. 13.20 Symmetry-breaking excitonic instabilities in a deformable lattice, with possible condensation of excitons to form a new phase with charge-density wave.⁴⁸

type. Namely, an electron is optically excited from Pt^{2+} to a neighboring Pt^{4+} (both becoming Pt^{3+}), whereby the surrounding X^- ions are brought into vibration along the chain axis due to the imbalanced forces caused by the charge transfer, relaxing towards their new equilibrium positions which amounts to the local formation of a hypothetical Pt^{3+} lattice. The situation can also be depicted as in Fig. 13.20 in which a hypothetical exciton in the *false* ground state (surrounded by a broken line) of the hypothetical structure (all Pt^{3+}) is supposed to be self-trapped, and self-decomposed into a pair of ions (Pt^{2+} and Pt^{4+}). Because the latter (relaxed excited state) has negative energy, the ion pairs are spontaneously generated at all sites without optical excitation, which is nothing other than the *real* structure shown at lower left. Note that the hypothetical exciton in the hypothetical material is not that real exciton observed in the real material which is not shown in this figure. The situation is one example of that shown in Fig. 13.1.

Another interesting observation was made on the same materials by Suemoto and collaborators.⁵⁵ The time-resolved emission spectra of an exciton in the femto second (fs) regime revealed the process of self-trapping of the exciton which can be analyzed in terms of an oscillating emission peak, as is described in Figs. 4.3(a) and 11.6(a) although these figures describe the relaxation of a localized excitation instead of the self-trapping of an exciton. The observed period of oscillation is somewhat smaller than the above-mentioned spacing of the Raman lines due to the participation of acoustic modes in the interaction mode.

More recent study of the same material by Sugita *et al.*⁵⁶ with the use of 5-fs resolution has revealed (1) the existence of two periods of the oscillation attributed to the ground and excited electronic states and (2) a delay time for the free exciton to be self-trapped which is interpreted to be the time spent at the flat part of the potential (without barrier in this one-dimensional system).

For quantitative analysis of these spectra, see the theoretical study by Kayanuma.^{57,58} [The effective period of the relaxational oscillation is about $3/5$ times ($2\pi/\omega$) in Fig. 1 of the former paper according to a private communication by the author.] A theoretical study of the time-resolved spectra of the self-trapping process of an exciton in a one-dimensional system (instead of the relaxation of an excited localized electron) has been presented by Ishida⁵⁹ using the molecular dynamics approach and by Tanaka and Kayanuma⁶⁰ using the time-dependent variational method. It is interesting to note that a delay time has been found in exciton self-trapping in a one-dimensional system in these theories, which may be related to observations.

In contrast to the above-mentioned compounds and those with Pd replacing Pt which have CDW ground states, those with Ni replacing Pt have an SDW ground state with halide ions situated equidistant from neighboring Ni^{3+} ions because of the large values of U of the Ni-3d state (in accordance with the phase diagram at the top of Fig. 13.17), as has been clarified by extensive studies by Yamashita *et al.*⁶¹ They estimated U to be as large as 5 eV, while the absorption edge at around 1.5 eV was assigned to the charge transfer transition from halide ion to Ni^{3+} (upper Hubbard band). Namely, these compounds are charge transfer insulators instead of a Mott insulator, the lower Hubbard band being below the halide band. The competition between the SDW and CDW states has been studied in great detail on mixed crystals of Ni and Pd compounds with variable compositions,⁶¹ while an elaborate study of photogenerated solitons, polarons and excitons with their dynamics has been performed by Okamoto *et al.*⁶²

13.6.4 Three-dimensional system with N sites and N_e electrons

A mean field theory for $T-U-S$ problems has been worked out by Shinozuka,⁶³ inclusive of the general case $N_e \neq N$ ($\sum_{\ell} n_{\ell}$, the total number of electrons, commutes with the total Hamiltonian and can be specified as N_e independently of N). Introduce $n \equiv N_e/2N$, the average number of electrons per site and per spin state. The system is now supposed to be an alloy consisting of four kinds of site, specified accordingly as

$$\begin{aligned} (m = 0) \quad & \langle n_{\ell\uparrow} \rangle, \langle n_{\ell\downarrow} \rangle \leq n \\ (m = \uparrow) \quad & \langle n_{\ell\uparrow} \rangle \geq n \geq \langle n_{\ell\downarrow} \rangle \end{aligned}$$

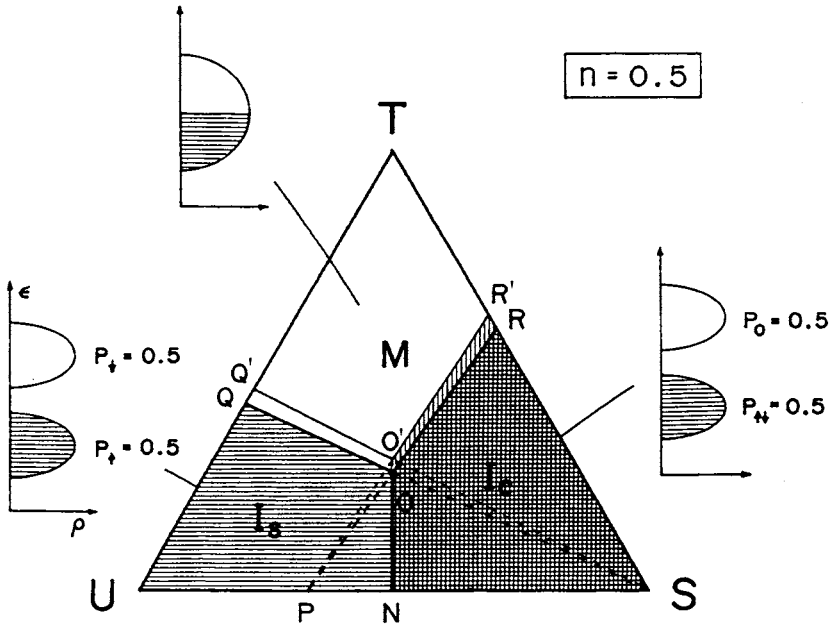


Fig. 13.21 Phase diagram for a half-filled electron case. Horizontal hatching represents insulating phase and vertical hatching represents phase with structural change. The density of states for \uparrow -spin electron is schematically illustrated for each phase (hatching represents occupied states). Due to Shinozuka.⁶³

$$\begin{aligned} (m = \downarrow) \quad & \langle n_{\ell\downarrow} \rangle \geq n \geq \langle n_{\ell\uparrow} \rangle \\ (m = \uparrow\downarrow) \quad & \langle n_{\ell\uparrow} \rangle, \langle n_{\ell\downarrow} \rangle \geq n, \end{aligned}$$

which are spatially distributed at random with the fractional concentration P_m . The electronic states in this alloy are solved with the use of the coherent potential approximation, by considering electrons in the semi-spherical density of states of the non-perturbed band (T) to be scattered by electron-electron (U) and electron-phonon (S) interactions at each site characterized by m . With the thus-obtained spin-dependent density of states to be occupied according to the Fermi statistics, the local density of spin and charge, P_m and the lattice distortion are determined self-consistently. The resulting phase diagram is shown in Figs. 13.21 and 13.22 for the cases of $n = 0.5$ ($N_e = N$) and 0.25 ($N_e = N/2$), respectively. The metallic phase M , the insulating phases I_s and I_c consisting of sites with nonvanishing spin or charge, appear around the T -, U - and S -vertex, respectively, and small transition regions appear which make M - I_s and M - I_c transitions to be of the second order. In the case of $n = 0.25$, there appears also the M_s -region, metallic but with sites of nonvanishing spin.

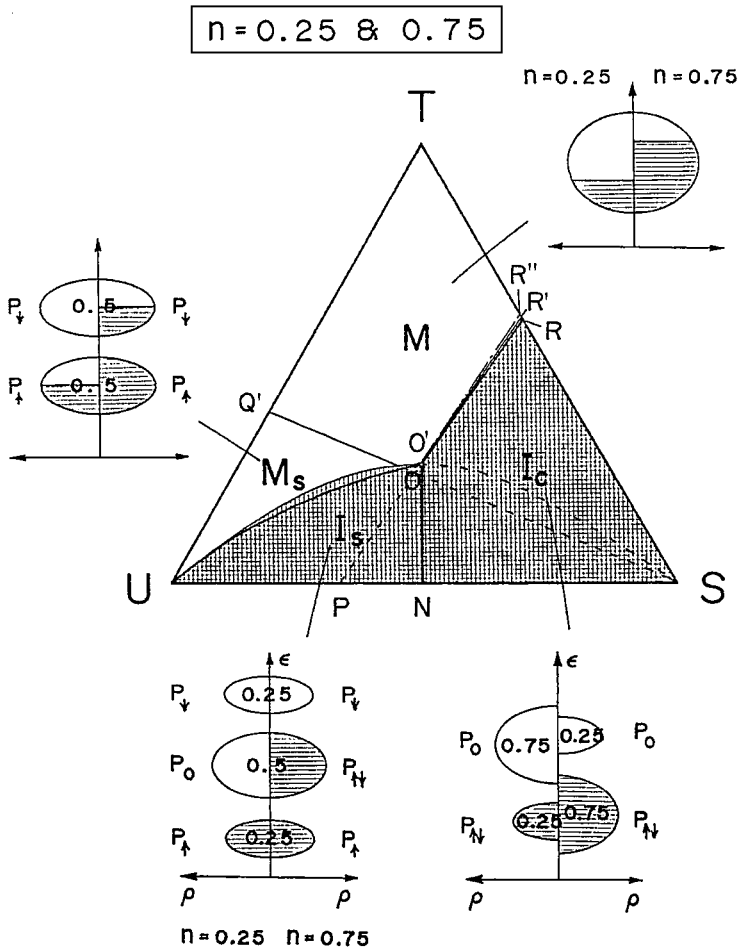


Fig. 13.22 Phase diagram for quarter- and three-quarters-filled cases. Horizontal hatching represents insulating phase and vertical hatching represents phase with structural change. The density of states for \uparrow -spin electron is schematically illustrated for each phase of a quarter- (*left*) and three-quarters- (*right*) filled cases. Hatching represents occupied states. Due to Shinozuka.⁶³

13.6.5 Tetrahedral 4-site system with variable occupancy N_e

This is an interesting model for a point defect in tetrahedrally coordinated covalent semiconductors such as column IV elements and III–V compounds, with sites representing dangling bonds for a vacancy case and bonding or antibonding orbitals for a substitutional impurity. The optical excitation and ionization give rise to changes in structural symmetry and occupancy, which provide plenty of chances to compare theory with observations. An extensive study of the ground and excited states has been performed by El-Maghraby and Shinozuka,⁶⁴ who found stable

configurations of distortion with symmetries T_d , C_{3v} , C_{2v} , C_s or C depending on T , U , S as well as on the electron number $N_e (= 0-8)$. An important difference from the ring systems is the asymmetry between the electron and the hole originating from the asymmetry of the unperturbed ($U = S = 0$) energy level structure: $E = -3T$ and $+T$ (triply degenerate). For instance, the off-center instability of a substitutional deep impurity is found to be more favorable for acceptor than for donor.

13.7 Extrinsic self-trapping, and shallow–deep bistability of donor state

We will be concerned here with the self-trapping of an electronic carrier(s) (electron, hole or exciton) at a point defect of the lattice such as a substitutional impurity, an antisite atom, vacancy, etc. It includes the situation in which phonon-induced localization of an exciton takes place *at an isoelectronic impurity but not in the host lattice* (*extrinsic self-trapping* which was mentioned briefly at the end of Section 9.5), as well as the situation in which a donor in semiconductors normally with *weak* electron–phonon coupling achieves a *deep state* with *strong lattice relaxation* by *capturing another electron* (D^- state with negative U , the so-called DX center) in addition to the *shallow state* (D^0). In either case, there are *stable* and *metastable* states separated by a potential barrier of the adiabatic potential. These two situations will be studied in the same context in this section although historically they were found and studied independent of one another.

It is conventional to refer to the lattice configuration in the delocalized (in the case of a neutral defect) or shallow (in the case of a charged defect such as a donor) state of the carrier(s) as the *undistorted* lattice and to consider that of the localized or deep state (with greater electronic binding energy and with stronger localization) to represent *strong lattice relaxation*. One finds more of an *anomaly* in the localized or deep state from the conventional viewpoint that the delocalized or shallow state is the normal state of the carrier. For example, the luminescence from the extrinsic self-trapped exciton is a *broad band* with *large Stokes shift* in contrast to a sharp resonant emission from a delocalized exciton, and a deep donor (D^-) has an optically accessible conducting state (a shallow donor D^0 plus an electron in the conduction band) which is *long-lived* due to the potential barrier separating it from the stable deep state; one can thus observe a *persistent photocurrent* even a day after optical excitation.

We will begin with a qualitative study of the condition for the appearance of this bistability, in particular the general trend dictating how the relative stability of the deep and shallow states depends on the physical parameters characterizing the carrier, the defect and the host crystal.

In the above description of what has so far been found experimentally, one may well wonder why in the donor case two electrons can cause a lattice distortion

much greater than twice that caused by a single electron, creating a shallow–deep bistability with a potential barrier. The answer to this and related questions will be obtained through the following qualitative argument.

Let us start with the adiabatic energy functional (9.5.1) in the continuum model for the interacting system of an electron and the phonon field, and introduce a point defect at $\mathbf{r} = 0$ with charge Ze ($Z = 0$ for an isoelectronic impurity and $Z = 1$ for a donor atom) and an onsite potential v_d (which depends upon the combination of the guest atom and the host lattice). One can then write the energy functional as⁶⁵

$$\begin{aligned}
 E[\psi, \Delta, \Phi] = & \int d\mathbf{r} \psi(\mathbf{r}) \left[-(\hbar^2/2m_e) \nabla^2 + v_d a_0^3 \delta(\mathbf{r}) - Ze/4\pi\epsilon_e r \right] \psi(\mathbf{r}) \\
 & + \int d\mathbf{r} \{ \Xi[\psi(\mathbf{r})]^2 + \Xi_d \delta(\mathbf{r}) \} \Delta(\mathbf{r}) \\
 & + \int d\mathbf{r} \{ (-e)[\psi(\mathbf{r})]^2 + Ze \delta(\mathbf{r}) \} \Phi(\mathbf{r}) \\
 & + \int d\mathbf{r} (C/2) [\Delta(\mathbf{r})]^2 + \int d\mathbf{r} (\epsilon/2) [\nabla \Phi(\mathbf{r})]^2, \tag{13.7.1}
 \end{aligned}$$

where the second terms in the curly brackets of the second and third integrands represent the interactions of the point defect with the dilation field and the electrostatic field, respectively. Here a_0 is the interatomic distance in the host lattice, and ϵ is a constant defined by $\epsilon^{-1} \equiv \epsilon_e^{-1} - \epsilon_s^{-1}$ with ϵ_e and ϵ_s being electronic and static dielectric constants, respectively (see the explanation between eqs. (9.5.1) and (9.5.2)).

By minimizing (13.7.1) with respect to ψ , one would obtain the adiabatic potential $W[\Delta, \Phi]$ for the lowest electronic state in the ∞ dimensional configuration coordinate space ($\Delta(\mathbf{r})$, $\Phi(\mathbf{r})$). Since this can be done only in a restricted way (see the argument associated with (9.5.9–9.5.17)), we will confine ourselves to the extrema of the adiabatic potential $W[\Delta, \Phi]$ as was done in the earlier part of Section 9.5. Namely, as the first step, we extremize (13.7.1) with respect to Δ and Φ , and obtain the “field-source” relations:

$$-C \Delta(\mathbf{r}) = \Xi[\psi(\mathbf{r})]^2 + \Xi_d \delta(\mathbf{r}), \tag{13.7.2}$$

$$\epsilon \nabla^2 \Phi(\mathbf{r}) = (-e)[\psi(\mathbf{r})]^2 + Ze \delta(\mathbf{r}). \tag{13.7.3}$$

One can imagine with (13.7.2) that a substitutional impurity with radius smaller or larger than the host atom it replaces has positive or negative Ξ_d , respectively, and correspondingly that an interstitial atom and a vacancy have negative and positive Ξ_d , respectively. Putting (13.7.2 and 13.7.3) into (13.7.1) and removing the self-energies of the defect itself in the phonon field ($\propto \Xi_d^2, \propto (Ze)^2$) which do not

depend on the electronic state, we obtain the functional

$$E[\psi] = \int d\mathbf{r} \psi(\mathbf{r}) [-(\hbar^2/2m_e)\nabla^2 + (v_d a_0^3 - \Xi \Xi_d/C)\delta(\mathbf{r}) - Ze^2/4\pi\epsilon_r] \psi(\mathbf{r}) \\ - (1/2) \iint d\mathbf{r} d\mathbf{r}' [\psi(\mathbf{r})]^2 [(\Xi^2/C)\delta(\mathbf{r} - \mathbf{r}') + e^2/4\pi\epsilon |\mathbf{r} - \mathbf{r}'|] [\psi(\mathbf{r}')]^2. \quad (13.7.4)$$

As the second step, we extremize $E[\psi]$ with respect to ψ to obtain the extrema of the adiabatic potential.

The difference of (13.7.4) from (9.5.2) is in the second and third terms in [...] of the first integral which represent the potential originating from the defect. It is to be noted that the product of the first and the second terms on the r. h. s. of (13.7.2) contributes the short-range impurity potential $(-\Xi \Xi_d/C)\delta(\mathbf{r})$, the defect-induced deformation potential. The corresponding quantity from (13.7.3) contributes the long-range screening potential $Ze^2/4\pi\epsilon_r$ which screens the impurity Coulomb potential $-Ze^2/4\pi\epsilon_r$ into $-Ze^2/4\pi\epsilon_s r$. Proceeding in the same way as in Section 9.5 with the use of the same trial function $\psi(r)$, we obtain the adiabatic energy of this system, consisting of *one* electron (suffixed or superfixed with 1) in the phonon (with L) and defect (with d) fields, in a form similar to (9.5.5),

$$W_{Z,1}(a) \equiv E_{Z,1}(\lambda) = B\lambda^2 - E_s^{Z,1}\lambda^3 - E_\ell^{Z,1}\lambda, \quad (13.7.5)$$

$$\text{with } E_s^{Z,1} \equiv E_s^L + E_s^d, \quad E_\ell^{Z,1} \equiv E_\ell^L + E_\ell^d, \quad (13.7.6)$$

$$E_s^L \equiv \Xi^2/2Ca_0^3, \quad E_s^d \equiv 2^{3/2}(-v_d + \Xi \Xi_d/Ca_0^3), \quad (13.7.7)$$

$$E_\ell^L \equiv e^2/4\pi\epsilon a_0, \quad E_\ell^d \equiv 2^{3/2}Ze^2/4\pi\epsilon_s a_0. \quad (13.7.8)$$

The short- (s) and long- (ℓ) range parts of the defect potential E^Z play the same roles as the corresponding parts of the electron–phonon interaction E^L . In fact, the phase diagram for the stable and metastable states is given by Fig. 13.23, which is identical to Fig. 9.8 except that the abscissa and ordinate represent the redefined short- and long-range coupling constants:

$$g_s^{Z,1} \equiv E_s^{Z,1}/B, \quad g_\ell^{Z,1} \equiv E_\ell^{Z,1}/B, \quad (13.7.9)$$

respectively. For clarity, the adiabatic potential in the case of a *charged* defect (Z : positive integer) is shown schematically in each region of the phase diagram, with shallow (F)–deep (S) bistability in the regions F(S) and S(F) where the state in the parenthesis is the metastable state. As is obvious from the adiabatic potential in the S(F) region, the optical ionization energy (distance from the deep state S vertically to the ionization continuum) is much larger than the thermal ionization energy (distance non-vertically to the minimum point of the ionization continuum). What is really deep is the optical (purely electronic) binding energy, not the thermal one.

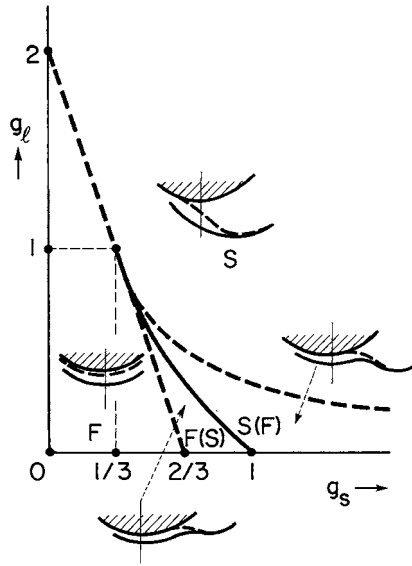


Fig. 13.23 Phase diagram for the stable (metastable when inside the parenthesis) state of an electron subject to short-range (g_s) and long-range (g_l) interactions with the defect and the phonon fields. F and S mean a shallow (delocalized for a $Z = 0$) state and a deep state, respectively. The adiabatic potential is also shown schematically for each region, where the lowest line should be replaced by the broken line when $Z = 0$.⁶⁵

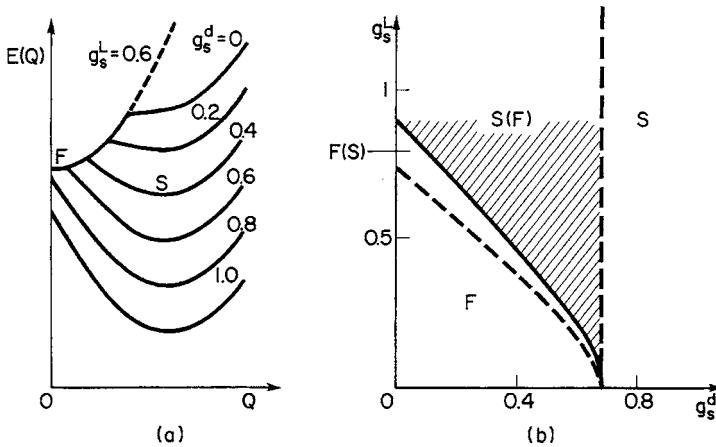


Fig. 13.24 (a) Adiabatic potential of a carrier coupled to a short-range attractive potential and a short-range phonon field through g_s^d and g_s^L , respectively. (b) Corresponding phase diagram (for “extrinsic self-trapping”) on the (g_s^d, g_s^L) -plane.⁶⁶

The adiabatic potential for an exciton ($E_\ell^L = 0$) extrinsically self-trapped at an isoelectronic impurity ($E_\ell^d = 0$) is given in Fig. 13.24, together with the corresponding phase diagram on the coordinate plane of the short-range electron–phonon ($g_s^L \equiv E_s^L/B$, ordinate) and electron–defect ($g_s^d \equiv E_s^d/B$, abscissa) couplings.⁶⁶

Extrinsic self-trapping of an exciton was found in TlCl:Br by Takahei and Kobayashi⁶⁷ and also in ZnSe:Te by Lee, Mysyrowicz and Nurmikko⁶⁸ through the *Stokes-shifted broad* emission band which is *absent in the pure* crystal. Because of the Franck–Condon principle, the effect of lattice relaxation does not manifest itself in the absorption spectra even when extrinsic self-trapping takes place. The dependence of the emission band on the impurity concentration is quite intriguing, showing some indication of impurity percolation.

One can calculate the corresponding expressions for the system of two electrons in the spin singlet state with identical orbital $\psi(\mathbf{r}_j)$ ($j = 1, 2$). One has to multiply by 2 the one-electron terms such as the first terms on the r.h.s. of (13.7.2) and (13.7.3) and the first integral in (13.7.4) and multiply by 4 the self-interactions (the second integral in (13.7.4)), and add the electron–electron Coulomb energy

$$U = (e^2/4\pi\epsilon_e) \iint d\mathbf{r}_1 d\mathbf{r}_2 \psi(\mathbf{r}_1)^2 |\mathbf{r}_1 - \mathbf{r}_2|^{-1} [\psi(\mathbf{r}_2)]^2. \quad (13.7.10)$$

The adiabatic energy is then given by

$$W_{Z,2}(\lambda) = 2[B\lambda^2 - E_s^{Z,2}\lambda^3 - E_\ell^{Z,2}\lambda], \quad (13.7.11)$$

$$\text{with } E_s^{Z,2} \equiv 2E_s^L + E_s^d = E_s^{Z,1} + E_s^L, \quad (13.7.12)$$

$$E_\ell^{Z,2} \equiv 2E_\ell^L + E_\ell^d - U/2 = E_\ell^{Z,1} - E^C, \quad (13.7.13)$$

$$E^C \equiv e^2/4\pi\epsilon_s a_0 = (\epsilon/\epsilon_s)E_\ell^L. \quad (13.7.14)$$

Note that the second term on the far right hand side of (13.7.13) is the sum of $E_\ell^L = e^2/4\pi\epsilon_e a_0$ and $-U/2 = -e^2/4\pi\epsilon_e a_0$. The phase diagram for the stable and metastable states is again given by Fig. 13.23, with the abscissa and the ordinate representing

$$g_s^{Z,2} \equiv E_s^{Z,2}/B \text{ and } g_\ell^{Z,2} \equiv E_\ell^{Z,2}/B, \quad (13.7.15)$$

respectively.

It should be noted that we are assuming that the two electrons always occupy the same orbital $\psi(\mathbf{r})$. Hence, the bistability studied with the above analysis and in terms of the phase diagram of Fig. 13.23 is between the deep and shallow states of D^- in the case of $Z = 1$ to which we shall confine ourselves for the moment. The discrete line in its inset represents the adiabatic potential of the D^- state while the hatched continuum that of the D^0 state plus an electron in the conduction band. The orbital radius in the shallow D^- state (denoted by (ii) to distinguish it from the states (i) and (iii) mentioned below) is larger than in the shallow D^0 state, due to the electron–electron Coulomb repulsion in the former. Other types of the shallow state of the two-electron system come into play in realistic situations: (i) the shallow D^0 state plus one electron in the *conduction band* (relevant to the description of the

persistent current), and (iii) the shallow D^0 state plus another electron in *another shallow* D^0 state at a distance. (The energy difference U^* between the deep state of D^- and the state (iii) can be negative when the Coulomb repulsion U is exceeded by the phonon-mediated attraction between the electrons, a situation which is called “negative U ”). For the energy analysis of these three *shallow* states, it is more appropriate to make use of the hydrogenic wave function $\psi = (\alpha^3/\pi)^{1/2} \exp(-\alpha r)$ instead of the Gaussian function used in the above and in Section 9.5. Minimizing the energy corresponding to eq. (13.7.9) with respect to α by neglecting the α^3 term, one obtains the energies of the above states as:

$$(i) W_{1,1} + W_{0,1} = -\varepsilon_d(1 + \gamma_p)^2 - \varepsilon_d\gamma_p^2, \quad (13.7.16)$$

$$(ii) W_{1,2} = -2\varepsilon_d(1 - \gamma + \gamma_p)^2, \quad (13.7.17)$$

$$(iii) 2W_{1,1} = -2\varepsilon_d(1 + \gamma_p)^2, \quad (13.7.18)$$

where

$$\gamma \equiv (1/2) \iint d\mathbf{r} d\mathbf{r}' [\psi(\mathbf{r})]^2 |\mathbf{r} - \mathbf{r}'|^{-1} [\psi(\mathbf{r}')]^2 / \int d\mathbf{r} [\psi(\mathbf{r})^2] r^{-1} = 5.16, \quad (13.7.19)$$

$$\gamma_p \equiv \gamma(\epsilon_s/\epsilon) \quad (13.7.20)$$

are the (negative) effect of electron–electron Coulomb repulsion and the polaron effect on the reciprocal radius α of the donor electron, and

$$\varepsilon_d \equiv (m_e/m)(\epsilon_0/\epsilon_s)^2 Ry \quad (13.7.21)$$

is the donor electron binding energy in the effective mass approximation with screening constant ϵ_s . The energies (13.7.16) to (13.7.18) are in the decreasing order: (i) > (ii) > (iii) provided $\epsilon_s/\epsilon_e > 1.233$. The last condition, needed only for the inequality (i) > (ii), should in fact be removed since an improved calculation with electron correlation taken into account gives (i) > (ii) even for nonpolar crystals ($\epsilon_s/\epsilon_e = 1$). Namely, a singly-charged defect can capture two electrons ((i) > (ii)) while two singly-charged defects prefer capturing one electron each ((ii) > (iii)). The ionization energies of D^- and D_0 in unit of ε_d are given by (i) – (ii) = $(1 - 4\gamma + 2\gamma^2) + (2 - 4\gamma)\gamma_p$ and (i) – (iii) = $1 + 2\gamma_p$, respectively. Since the short-range electron–phonon interaction is negligibly small for these shallow ($\alpha \ll 1$) states, there are no potential barriers between any two of them in contrast to that between the shallow and deep D^- states.

Let us now ask whether two self-trapped charge carriers (there are many examples of self-trapped holes but rather few of self-trapped electrons) can form a bound pair in a defect-free crystal. Dropping the defect terms in (13.7.5) to (13.7.8) and

(13.7.11) to (13.7.14), one obtains the binding energy as

$$2W_{0.1}(\gamma = 1) - W_{0.2}(\gamma = 1) = 2B[g_\ell^{0.2} - g_\ell^{0.1} + g_s^{0.2} - g_s^{0.1}] \\ = -2e^2/(4\pi\epsilon_s a_0) + 2\Xi^2/(2Ca_0^3). \quad (13.7.22)$$

Namely, the bound state can be formed when the phonon-mediated attraction (the second term) exceeds the Coulomb repulsion (the first term). That might well be possible for positive holes in alkali halides. To be more realistic, however, the effective screening constant for two holes at atomic distance is expected to be significantly smaller than the static dielectric constant ϵ_s of bulk matter, making the binding more difficult.

As is seen from Figs. 9.8 and 13.23, the increase in the long-range coupling constant g_ℓ has the effect of *shrinking* the region (seen along the horizontal line with constant g_ℓ) of *bistability* (F(S) plus S(F)) although it helps to induce larger lattice polarization and greater binding energy. Moreover, it should be noted that the adiabatic bistability is increasingly blurred towards the top cusp (see Fig. 13.23) of the bistability region by the quantum-mechanical tunneling of atomic motions across the potential barrier (the barrier height tends to zero as one approaches the cusp).

Note that the existence of a donor gives rise to a significant increase of g_ℓ according to (13.7.6) and (13.7.7), but that the participation of the second electron decreases it again to some extent according to (13.7.13) ($g_\ell^L < g_\ell^{Z,2} < g_\ell^{Z,1}$). Namely, the donor charge is unfavorable for the bistability, but the screening by another electron mitigates that, making the donor closer to a neutral center, and favoring the bistability (but probably not as much as in the host lattice). In contrast, the short-range coupling constant g_s is certainly increased by the participation of the second electron according to (13.7.10), possibly causing a large lattice distortion (as the representative point (g_s, g_ℓ) traverses the F–S discontinuity line). The short-range defect term (the second term on the r.h.s. of (13.7.6)) can work, depending on its sign, positively or negatively for the stabilization of the deep state of the two-electron system.

Let us consider an electron or hole in a typical semiconductor which is normally free ($(Z, n) = (0, 1)$ is within the region F in Fig. 13.25), and study how their states change under the existence of an (attractive) charged defect ($Z = 1$) and/or another carrier ($n = 2$) of the same species. A typical situation expected is shown schematically in Fig. 13.25. For comparison, a situation in alkali halides is also shown schematically where (1, 1) and (1, 2) represent the F center (see Sections 4.7 and 4.8) and an F' center, respectively. As a general trend, it is well conceivable that the vector $\bar{g} \equiv (g_s, g_\ell)$ on the phase diagram of Fig. 13.25 for one electron in the conduction band ($\bar{g}^{0.1}$) as well as at the donor ($\bar{g}^{1.1}$) is well inside the F-region, while for two electrons captured by the donor ($\bar{g}^{1.2}$) it is situated in the bistable

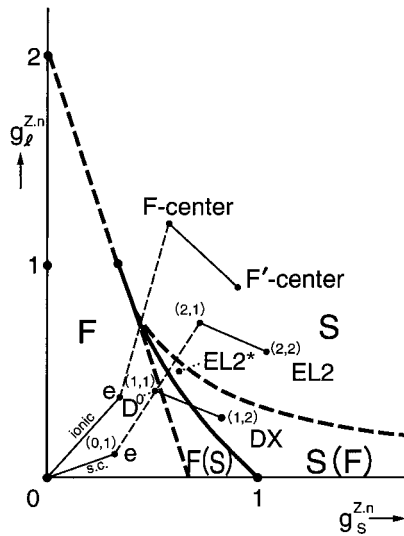


Fig. 13.25 Typical location of the states (Z, n) of n electrons at the defect with charge Ze in weakly polar semiconductors (s.c.) and in ionic crystals on the $(g_s^{Z,n}, g_\ell^{Z,n})$ -plane of the phase diagram.

S(F)-region (although the bistability of one electron at a donor cannot be excluded). Combining this with the condition for the binding of two charged carriers mentioned before, we can conclude as follows.

A large lattice distortion can be caused by the cooperation of two carriers even when it is not caused by a single carrier as shown in Fig. 13.25 where the shallow-deep discontinuity line divides (1, 1) and (1, 2) on its opposite sides, provided that they are bound together. Because of their Coulomb repulsion, however, the phonon-mediated attraction by itself is not enough but the assistance of the attractive charged defect is needed to bind them together, in spite of the fact that the charged defect by itself works unfavorably for the bistability.

The general trend of all that has been observed – the presence of an exciton ($g_\ell^L = 0$ due to its neutrality) bistability (intrinsic and extrinsic self-trapping with potential barrier), few known cases of (a single) charge carrier ($g_\ell^L \neq 0$) bistability in pure and impure polar crystals (neither the self-trapped hole nor the F-center electron in alkali halides has a potential barrier from the delocalized state), and a number of defects with two electrons showing bistability which will be described in more detail later, are all consistent with the above conclusions. A bound pair of self-trapped holes has not yet been found even in alkali halides as is doubted above, but it may be found extrinsically near a negatively charged defect due to the reason mentioned above.

While the continuum model can describe the general trend of the conditions favorable for shallow–deep bistability, it cannot predict the microscopic details of

the strongly localized state. In fact, the microscopic structures of localized carrier(s) with bistability which have been revealed by a great number of experimental studies and first-principle calculations during the last twenty years are much more intriguing and versatile than can be imagined within the continuum model. Referring to Ref. [69] for a brief historical survey of these studies, we will describe here the microscopic structures so far revealed of three typical defects with bistability, the DX center and the EL2 center in covalent crystals and the DX-like center in ionic crystals, and reconsider them in the light of what has been discussed above and in the preceding sections.

The DX center is the negatively charged donor atom of a column IV (or VI) element substituting the column III (or V) site of a III–V compound semiconductor, such as GaAs:Si (or S). As shown in Fig. 13.26(a) or (c),⁷⁰ all host and guest atoms are covalently bonded to their four nearest neighbors when the donor is ionized (D^+) or neutral (D^0). While an electron captured by the D^+ center is in the shallow effective mass state with large radius, the capture of a second electron results in

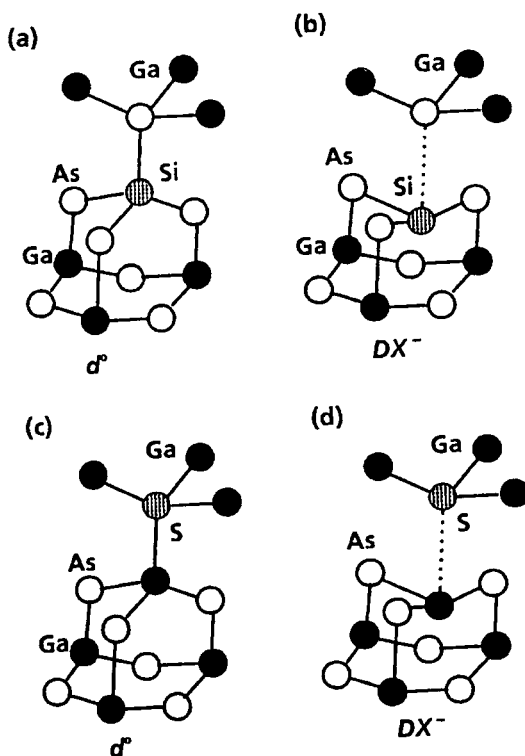


Fig. 13.26 Schematic views of the normal substitutional sites (*left*) and the broken bond configurations (*right*) giving rise to DX centers in Si- (upper) and S- (lower) doped $Al_xGa_{1-x}As$ alloys. Due to Chadi and Chang.⁷⁰

a non-centrosymmetrically distorted configuration (Fig. 13.26(b) or (d)), according to the self-consistent calculation of total energy by Chadi and Chang.⁷⁰ The stable atomic configuration of this “DX” center is such that the dopant and one of its neighbors have moved oppositely towards their respective interstitial sites. The bonding of these two central atoms to their neighbors has changed thereby from sp^3 (with four neighbors) to sp^2 (with three neighbors) configurations according to the conventional description of the covalent bond.

The valence- and conduction-band states of covalent semiconductors consists of bonding and antibonding orbitals, respectively, of each pair of neighboring atoms. Then the well-localized electrons in the deep donor state can nearly be considered to be in the antibonding orbital between the donor and one of its neighbors. To be more exact, the transition from the donor-centered state to the bond-centered state takes place with the pseudo Jahn-Teller displacements, as described by Dabrowski and Scheffler.⁷¹

According to the description proposed in Section 13.1, a covalent bond is formed by two *p-holes* in the spin singlet state occupying the *antibonding* orbitals. In the band model of covalent semiconductors, these *holes* are to *occupy* the states *above* the Fermi level which in n-type materials is around the bottom of the conduction band which is antibonding. The atomic and electronic configurations of the D^- state in Fig. 13.26 are then straightforwardly described as the *breaking* of the covalent bond between the donor atom and one of its neighbors by the *annihilation of the positive holes constituting that bond by the two electrons strongly localized* due to their cooperative effect. After bond breaking, the two electrons occupy the non-bonding orbitals of the two atoms on each end of the broken bond. This is a typical example in which the hole picture is preferable to the electron picture for the covalent bond.

While the bistable behaviors of the DX center can be described by associating it with the $(Z, n) = (1, 2)$ state and the $D^0(1, 1)$ plus $e(1, 1)$ states in Fig. 13.25, the situation is somewhat different in the EL2 center in GaAs, which is now generally agreed to be As_{Ga} , namely the antisite As atom.⁷² In the neutral state, this As atom occupying a Ga site has two electrons not participating in the covalent bond which are situated at around the middle of the bandgap. The center behaves as a deep donor in n-type materials and as a deep acceptor in p-type materials, with nearly equal ionization thresholds (~ 0.76 eV, one-half of the energy gap) as seen from photoluminescence excitation spectra.⁷³ This binding energy is too large to be described in the effective mass approximation for the divalent donor ($Z = 2$) even if the lattice relaxation is considered, and can only be explained by considering the local potential v_d . The main contribution to v_d comes from the fact that the donor potential given by $-Ze^2/4\pi\epsilon_s r$ at long distance is not so well screened at interatomic distance $r \sim a_0$. If one replaces the macroscopic screening constant

ϵ_s by an effective ϵ_{eff} which is supposed to be somewhere between ϵ_s and ϵ_0 , the difference $v_d = -(\epsilon_{\text{eff}}^{-1} - \epsilon_s^{-1})Ze^2/4\pi a_0$ can well be as large as several eV (with minus sign). Since this value significantly exceeds the bandgap (1.52 eV) of GaAs, the donor (and acceptor) level is expected to be around the midgap (in agreement with the observation) according to the calculation which takes into account the conduction band and the valence band (the donor level is equally repelled by the two bands). In the present effective mass approximation with only the conduction band taken into account as the basis, the effective v_d would have to be chosen so as to give the observed depth.

For our model calculation, let us take this effective short-range potential v_d to be proportional to Z , and neglect the second term of E_s^d in (13.7.7). Then both the short- and long-range parts of the defect potential are proportional to Z , and one can compare the situations in the divalent donor ((2, 1) and (2, 2)) to those in the monovalent donor ((1, 1) and (1, 2)) as shown schematically in Fig. 13.25. In this model case, the divalent donor is too deep to have bistability (S region) in contrast to the monovalent donor which is moderately deep to have a metastable state (S(F) region). This is in agreement with what is observed in the EL2 center. While its ground state is known to have a tetrahedrally coordinated structure without bistability, its optically-excited state is subject to Jahn-Teller distortion which relaxes to a configuration similar to that in the DX center according to Refs. [70, 71]. The relaxed excited state would correspond to an effective value Z^* between 1 and 2 (see the point EL2* in Fig. 13.25) since the excited electron with large radius will see the divalent donor core plus another electron left in the ground state as a monovalent donor. This off-center configuration is again considered to be a broken covalent bond between the antisite As with one of its neighboring As atoms due to the annihilation, by the excited electron, with one of the holes which was responsible for the covalent bond. Since the holes responsible for the covalent bonds are populated only above the Fermi level which is near the conduction-band bottom in the n-type material, only the excited electron can make the annihilation. The deep electrons in the ground state cannot find such holes to annihilate with in the nearby energy region, which explains the electronic and structural stability of the EL2 center.

It is obvious in these two examples that the electrons causing bond breaking are near the Fermi energy, and the resulting deep levels destroy the donor functionality of readily supplying electrons from shallow levels to the conduction band. In fact, discovering how to solve the doping difficulty encountered in some semiconductors seems to have been an important motivation for the study of deep levels, leading to the discovery of DX centers⁷⁴ and other bistable systems. The doping limit is more systematically studied by Zhang, Wei and Zunger.⁷⁵

A third example of bistability is the CdF₂ crystal doped with a column IIIB element such as Ga or In⁷⁶ which shows a bistability similar to the IIIB–VB

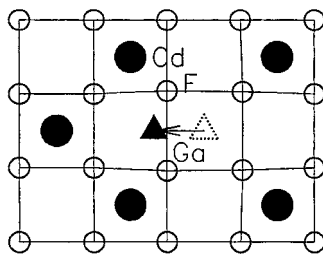


Fig. 13.27 Schematic structures of CdF_2 consisting of simple cubic lattice of fluorines with cations occupying the centers of alternate cubes. The substitutional position of donor core Ga^{3+} with one electron (D^0) and the displaced position of that with two electrons (D^-) are shown schematically by dotted and solid triangles, respectively. Due to Park and Chadi.⁷⁸

semiconductor doped with a column IVb element. In addition to its two-electron nature, the existence of an open volume in the deep donor state revealed by a positron annihilation study⁷⁷ gave convincing evidence of its similarity to the DX center. In fact, the first-principle calculation due to Park and Chadi⁷⁸ confirmed that the donor which normally substitutes the Cd site is displaced, by capturing another electron, into a neighboring empty cube of F ions, as shown in Fig. 13.27. In this ionic crystal, the interionic Coulomb interaction plays a dominant role and the Ga^+ ion, which is less repelled by Cd^{++} ions than Ga^{++} is, prefers this position by making use of the interaction between their d electrons according to their calculation, which is also endorsed by their result that this structure is not stable in the CaF_2 host without d electrons. Thus, the microscopic nature of the interatomic interactions which stabilize this center is rather different from, in spite of the similarity of its behavior to, the DX center in covalent semiconductors.

Concerning the shallow–deep bistability of the CdF_2/In center, one should note a remarkable difference of the optical spectra between shallow and deep states.⁷⁹ The absorption spectrum of the deep state rises at threshold reflecting its large optical depth of ~ 1.5 eV and has a width as large as 2 eV due to its strong electron–phonon coupling, while that of the shallow state is characterized by a much smaller optical depth and width, both being in accordance with the general argument in this section.

[Note added in proof]

In a book by Itoh and Stoneham,⁸⁰ recently published, some of the problems treated in this chapter are described more extensively and from a view-point somewhat different from ours, providing useful and complementary knowledge to interested readers.

14

Light, matter and life

(A scenario for the origin of life and its evolution)

14.1 Introduction

This chapter is of a different nature from the preceding chapters where more or less well-established principles of the spectroscopic studies of microscopic motions in condensed matter have been described, as is usual in textbooks and monographs. In this last chapter, however, the author is going to present his *personal viewpoint* on the role of solar radiation in creating life out of matter, sustaining living activities and driving the evolution of the living world – topics which have generally been considered to be beyond the scope of the physical sciences.

Obviously, life science is a vast interdisciplinary regime: the elucidation of life will need the cooperation of all areas of natural science. The majority of scientists working in any discipline will supposedly have a serious interest in the problem of life even when they do not mention that explicitly. The particular reason why the present author feels it necessary and useful to make remarks on this problem is that the various photochemical processes described in the preceding few chapters and the behaviors of charged particles in dielectrics described in other chapters have something to do with, and to shed light on, the role of solar radiation as an energy source and the role of water as a catalyst in living activities. Further investigation of these roles might contribute to constructing a bridge from our small corner of physical science, among many other such bridges under construction from different areas of science, towards elucidating the physical origin and evolution of life itself.

What the author can do as a theoretical physicist is present speculations based on the various empirical and experimental facts available. Nevertheless, there seem to exist many reasons based on well-known fundamental physical principles that lend support to these speculations. If unfortunately some of the speculations turn out to be wrong, the responsibility is entirely down to the author's misuse of his knowledge and never on the side of the basic principles themselves established by the cooperation of so many scientists. If fortunately some others prove to be true,

the merits are naturally due to the universality and ubiquity of the basic principles. Even in the case of failure, the author considers that the interested reader may be encouraged to find the reason for that failure and to seek for a better scenario. In this way one can draw closer to a deeper reality of Nature.

14.2 Energy source for creating living matter and sustaining living activities

Putting aside all mysteries of living organisms, the obvious energy source of living activities is the *chemical energy* stored in the living matter. In typical textbooks on biochemistry¹ one can find the energy content of food constituents which are also important constituents of living cells in general: carbohydrate contains 16, protein 17 and fat 37 kJ per gram dry weight.² It is this chemical energy that is released as heat or some other form of energy when the constituent is decomposed, by reaction with atmospheric oxygen, into inorganic matter such as water, carbon dioxide and nitrogen. This means that the same amount of energy is needed in order to synthesize the biochemical constituents from those inorganic materials, thereby releasing the oxygen. Chemical energy is defined here by referring to water, carbon dioxide, nitrogen and oxygen in their standard state, namely the *present environment of the earth*, as energy zero. It is *different from the conventional definition* in which H₂ and carbon are chosen instead of H₂O and CO₂, respectively, as reference materials.

It is more revealing, from a microscopic point of view, to estimate the chemical energy *per atom* of these constituents with the use of their known atomic compositions. One finds readily the figures: 1.3, 1.3 and 1.8 eV for carbohydrate, protein and fat, respectively,³ values which are (i) much greater than the thermal energy on the earth: $k_B T \sim 0.023$ eV, and (ii) of the same order as the typical energy of a photon from the sun: 1–3 eV, as shown in Fig. 14.1(a,b) Note (Fig. 14.1(a)) that the spectral peak of the intensity distribution depends on whether one takes the wavelength or the photon energy as the variable, as seen in eqs. (1.4.8) and (1.4.10). We take the latter – the photon-energy – as the abscissa in Fig. 14.1, instead of the wavelength conventionally used in spectroscopy, although the choice is not really important for our qualitative argument.

The more than an one order of magnitude difference mentioned above in (i) indicates why biomolecules behave apparently against the thermodynamics which govern the world of thermal energy. The biomolecules are, so to speak, far from thermal equilibrium, and it is hopelessly difficult to synthesize them in a merely thermal environment on the earth. Several tens times in energy mentioned in (i) above means, according to statistical mechanics, $\exp(-\text{several tens}) \cong 0$ times in probability!

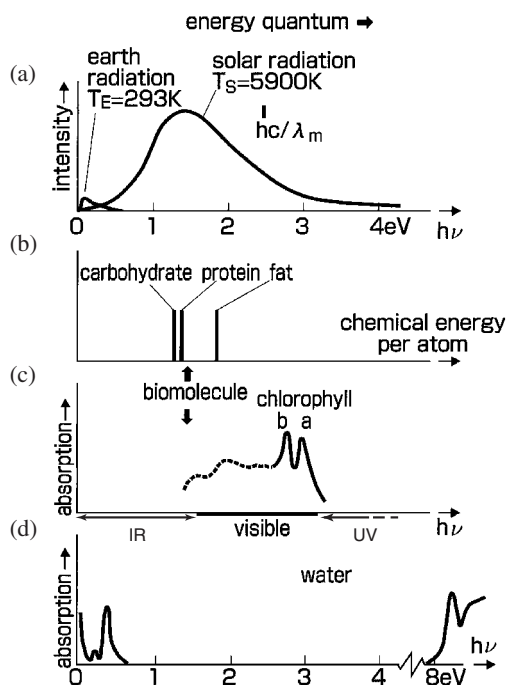


Fig. 14.1 Comparison of energy quanta relevant to living activities. (a) Spectral distributions of solar and earth radiation fields; (b) chemical energy *per atom* of typical biomolecules; (c) absorption spectra of photosynthetic pigments, with the highest-energy peaks of chlorophyll a and b shown by solid lines; and (d) absorption spectrum of water.

In order to synthesize a biomolecule from simple inorganic molecules by *successive addition of an atom or a molecule*, each elementary chemical reaction needs energy of the order of eV according to the above estimation. The fact (ii) mentioned above indicates that solar radiation is a good candidate for the promotion of successive synthesis of biomolecules. This is obvious also from the fact that the photosynthesis of biomolecules in plants under solar radiation has been and is now feeding all living organisms (inclusive of animals and humans who eat plants or other animals). It is of interest to find that chlorophylls have absorption spectra in the region of energy somewhat higher than the per atom chemical energy of biomolecules (see Fig. 14.1), which fact seems to contribute to the high efficiency of photosynthesis for the storage of chemical energy.

It is true that there are external energy sources with energy quanta *greater* than the solar photons, such as cosmic rays of various species, but they have a much smaller total energy flux than the solar radiation. "Solar radiation" as defined here includes only light – electromagnetic waves – from the sun, but does not include the other rays of high-energy particles. The electric discharge in the lightning which may have frequented the primordial earth has also been considered as a candidate for

Table 14.1 Available energy on the primitive earth.

Solar light	Energy, mW cm ⁻²	Energy yield, nmole joule ⁻¹
far UV < 200 nm	0.004	$1 \times \phi^a$
UV 200–300 nm	0.4	$2 \times \phi$
near UV 300–400 nm	9	$3 \times \phi$
Visible 400–800 nm	70	$6 \times \phi$
Electric discharge	0.0005	10
Radioactivity	0.0001	?
Volcanoes	0.00001	?
Shock waves, meteors, etc.	0.00001	100

The data are from Miller and Orgel (1973).

^a ϕ is the quantum yield of the photoreaction: moles of products/Einsteins or moles of photons absorbed.

Source: Mauzerall, D. (1992) Oceanic sunlight and origin of life, in *Encyclopedia of Earth System Science*, Vol. 3, pp. 445–453 (Academic Press, Inc.).

the energy source for the synthesis of biomolecules. However, lightning storms are nothing other than local and temporal meteorological phenomena caused by the solar radiation, and lightning cannot itself be the main source as viewed from the long history of the earth. The hydrothermal vents found in the deep bottom of the sea, an energy source of terrestrial origin in contradistinction from the above-mentioned sources, have also been studied seriously, partly because of their temperature as high as several hundreds of degrees kelvin. While the corresponding thermal energy is still much smaller than the energy per atom of the biochemical constituents mentioned above, a more serious objection against this origin is that the thermal source cannot by itself create the situation of population reversal as mentioned above. Speculations on an extraterrestrial origin of life on earth, such as the panspermia theory, do not help since they merely shift the origin of life to somewhere else in (or outside of) the universe without solving the question (how life was created there). Moreover, even if a small germ of a biomolecule of extraterrestrial origin did happen to reach this earth, it would have needed to be converted into a biopolymer here on the earth in order to have living activities, a task which itself is much more laborious than creating the small germ.

While there are many arguments favoring solar radiation as the source energy for living activities from the viewpoint of *total energy flux* as shown in Table 14.1 which is data due to Miller and Orgel,⁴ and Mauzerall,⁵ the above argument from the viewpoint of *an energy quantum supplied per an elementary photochemical process* lends *independent* support to solar radiation as the energy source for *creating life* as well as for sustaining living activities. Further support will be given towards the end of Section 14.3.

For this reason, we will study the following scenario for the origin of life on the earth mostly following Oparin and Haldane, the pioneers of the concept of chemical evolution (see Section 1–4 of Ref.[1]): there took place successive photochemical reactions in the prebiotic sea or ponds with simple inorganic (and possibly some organic) molecules dissolved therein under the solar radiation impinging on, or penetrating to some depth below, its surface. In our scenario, we will consider more seriously *how, to what extent and with what rate the photon energy absorbed* by an aqueous solution is *stored as chemical energy* of the synthesized molecules without re-emission of a photon and nonradiative de-excitation and decomposition, rather than specify the individual photochemical processes.

14.3 Conversion of radiation energy to chemical energy

Photoexcitation of a molecule, be it in atmosphere or in aqueous solution, gives rise to changes in its electronic charge distribution and interatomic bonding, destroying the balance of interatomic forces which kept the atomic positions as they were in the ground electronic state, as was described in various contexts in the preceding chapters. The atoms in that molecule and in the nearby medium (atmosphere or solvent) start to displace under the suddenly generated forces due to this imbalance, finally *relaxing* to their new equilibrium positions which depend upon the electronic state to which the molecule has been excited.

This new electronic and atomic configuration is called the *relaxed excited state* (RES), in contradistinction to the *Franck–Condon* (F.C.) state which refers to that before the atomic displacements (see the configuration coordinate (C.C.) model of Fig. 4.1). Here the one-dimensional abscissa represents the *interaction mode*, which is a linear combination of the displacements of all relevant atoms as projected along the direction of *steepest descent* of the excited-state adiabatic potential in *multi-dimensional* C.C. space (see Section 4.4 and Fig. 4.2). It is almost equivalent to what chemists call the *reaction* coordinate.

As for the ordinate, the energy difference between the F.C. state and the RES – the relaxation energy – is imparted as vibrational energy to other modes irrelevant to the electronic system of our concern, and in the case of a large molecule or a molecule in solution, that energy is spatially dispersed and hence irreversibly lost (dissipated) as heat. In this sense, the ordinate of RES in the C.C. model is to be understood as the free energy, $F = U - TS$, rather than the energy U itself. Note that the first and second terms of F correspond to the energy of the F.C. state and the energy lost as heat during relaxation, respectively. This interpretation of the C.C. model is of vital importance when we consider the chemical energy (essentially the same concept as the free energy) of molecules under the varying environment of the earth from era to era as we will do in Section 14.7. In connection with this,

the *negative entropy*, which Schrödinger once claimed to be the source of living activity, should be described in terms of a *large free energy*, the storing mechanism of which is our problem here. (The statistical mechanical definition of entropy as $S = k_B \ln W$ due to Boltzmann, where W is the number of microscopic configurations corresponding to the macroscopic state concerned, allows no negative entropy, the very point to which Schrödinger dared to ascribe the mystery of living matter.)

Let us resort to Fig. 13.5(b) which can also be applied to non-excitonic systems if the potential barrier between the F.C. state (F) and the RES (S) is ignored. In the cases α and β where the ratio of the relaxation energy E_R to the F.C. excitation energy E_0 is not too large (note that the figure is presented for variable E_0 (increasing in the order of U, γ, β, α) with fixed E_R), the relaxed system will emit a photon and return to the ground state (from S vertically down to the point on the adiabatic potential α or β although the line is not shown) within a time of the order of 10^{-7} s (see eq. (3.6.7)). The energy difference between the absorbed and emitted photons (the Stokes shift) is dissipated as heat during the relaxations after optical excitation and de-excitation, *no energy being stored* in the relevant system after the optical cycle.

In the case γ of stronger relaxation (relative to E_0) where the adiabatic potential of the excited state intersects that of the ground state before reaching the RES as also shown in the left part of Fig. 14.2, the photoexcited system will stay in the RES (M_1 in Fig. 14.2) much longer, due to the absence of a radiative de-excitation channel therefrom, than in the cases α and β . The only channel of de-excitation is a nonradiative process which is much less probable than the radiative process

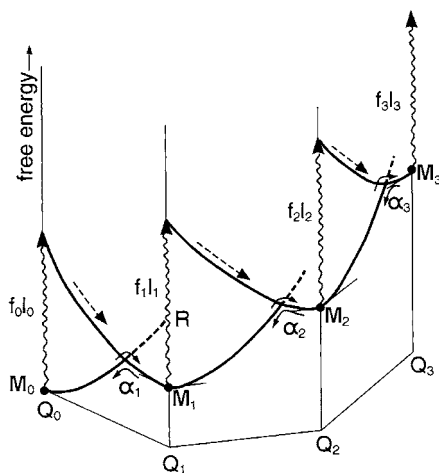


Fig. 14.2 Multi-dimensional configuration coordinate (Q_n) model with the adiabatic potentials (solid lines) for multistep photoexcitations (vertical wavy arrows), relaxations (along broken arrows) and nonradiative transitions (along solid arrows). As for 'R' in the figure, see the description towards the end of Section 14.7.

mentioned above as will be detailed later. Thus, the case γ , with its longer lifetime, is a situation most favorable for storage of the photoexcitation energy. Hereafter we will confine ourselves to the case γ .

If the interaction mode Q_0 – Q_1 of the C.C. model in Fig. 14.2 represents a *chemical reaction coordinate* (as in the case of Fig. 13.2 for excimer formation, which however belongs to the case β , not the γ of our concern), the RES (M_1) of our γ -type represents the *ground state of a new molecule* formed by a photochemical reaction in which the molecule M_0 is combined with another atom or molecule in the surrounding medium – solvent or atmosphere. (Note that the reaction coordinate $Q_0 \rightarrow Q_1$ corresponds to the decreasing distance of these particles from the central molecule M_0 as in the case of Fig. 13.2 for excimer formation.) The result is the storage of a part of the photexcitation energy as chemical energy $F = \text{length}(Q_1 M_1)$ of this new molecule M_1 . If a second photon hits this new molecule during its lifetime, there may take place a second photochemical process in which the molecule M_1 combines with another atom or molecule in the medium to form a molecule M_2 , as shown in the middle part of Fig. 14.2, and so on.

What fraction of the photon energy is stored as chemical energy in each photochemical reaction? To answer this question, let us recapitulate the argument in Section 13.3. We put $B = 0$ for our problem of localized excitation (not exciton), and slightly generalize the argument by assuming the curvature of the excited-state adiabatic potential to be λ times that of the ground state in Fig. 13.5(b) in order to meet more realistic situations. Namely, we put the adiabatic potentials of the ground and excited electronic states as

$$W_g(Q) = Q^2/2, \quad W_e(Q) = E_0 - cQ + \lambda Q^2/2. \quad (14.3.1)$$

The minimum of $W_e(Q)$ is at $Q_m = c/\lambda$, giving $W_e(Q_m) = E_0 - E_R$ with relaxation energy $E_R = c^2/2\lambda$. In order that the intersection of $W_g(Q)$ and $W_e(Q)$ appears at $Q < Q_m$ – the condition to distinguish γ from β (see Fig. 13.6) – the inequality

$$E_R > [\lambda/(1 + \lambda)]E_0 \quad (14.3.2)$$

must be satisfied, indicating that the slope of the boundary between S_γ and S_β is now given by $\lambda/(1 + \lambda)$ instead of $1/2$. The fraction Φ of the stored energy is therefore limited by

$$\Phi \equiv (E_0 - E_R)/E_0 < (1 + \lambda)^{-1}. \quad (14.3.3)$$

The ratio λ of the curvatures is related with that of the vibrational frequencies ω_e and ω_g in the excited and ground states by

$$\lambda = (\omega_e/\omega_g)^2. \quad (14.3.4)$$

In the case of intramolecular excitation, it is generally supposed that ω_e is smaller than ω_g since the interatomic bond and hence the interatomic force constant is weakened by the electronic excitation. This is also endorsed by the observations, in many molecules, of vibrational quanta $\hbar\omega_e$ and $\hbar\omega_g$ which appear in the absorption and emission spectra, respectively (see, for example, Fig. 10.8 for a NO_2 molecule). The situation is supposed to be more varied in the case of a photochemical reaction since a new bond with more *distant* atoms or molecules is formed in the excited state and the parabolic approximation (14.3.1) is expected to be poorer. Thus, one can expect that about half of the absorbed photon energy *can* be stored on each step of photochemical reaction of the type γ even under the restriction (14.3.3). This fact is consistent with the fact that the spectral region of strong absorption of chlorophyll a and b, among photosynthetic pigments, is on the higher energy (by 1.5–2 times) side of the chemical energy per atom of typical biomolecules as shown in Fig. 14.1(b, c). This consistency, in addition to the coincidence of the chemical energy with the typical photon energy of solar radiation mentioned before, lends further support to the model of successive photochemical reactions which will be studied more quantitatively in the next section.

14.4 Kinetic equations for chemical evolution under solar radiation

With the known intensity of solar radiation on the earth, and with the conceivable lifetime of nonradiative de-excitation in the case γ , the chance of successive photochemical reactions is in general very small as may readily be supposed, but will turn out to have a *significant result* in view of the *fantastically long* estimated *time-span* of 4.6 billion years since the formation of the earth. The extremely *small chance for survival* of the photochemical products, M_1, M_2, \dots necessitates a very *severe selection* of photoexcited states (among others) with *long-lived* RES (M_i) for each step, in order that they can make a significant contribution to the storage of solar radiation energy in the form of biopolymers. Notwithstanding this, one can expect a gradual but inevitably indefinite increase of the number and type of higher polymers, some of which may *apparently spontaneously* display what are called living activities by making use of their stored chemical energy.

For simplicity, we will confine ourselves to that photoexcited state, among others, with the *longest-lived* RES of type γ for each step, and consider a series of successive channels of photochemical reaction of this type, as shown in Fig. 14.2. Most of the non-selected optical transitions at each step will result in de-excitation with predominant probability and, hence, with no contribution to energy storage. In each step of photochemical reaction, the molecule combines with another atom or molecule in the medium and thereby grows larger. This extremely simplified model of a series of single-channel events is not as restrictive as it may appear

at first glance; for example, one can include the event in which the first molecule is excited in the first step, the second molecule in the second step, and in the third step one of them is photoexcited and comes closer to combine with the other to form a new molecule. It should be noted in passing that any RES reached by *chemical reaction* (instead of mere relaxation) represents the *ground state* of a new molecule, so that the optical excitation therefrom does *not mean* the *secondary* excitation.

The kinetic equation for the selected channel of photochemical reaction is then given by³

$$\begin{aligned} dP_n/dt = & [f_{n-1}I_{n-1} + \alpha_n\gamma_{n-1}]P_{n-1} + \alpha_{n+1}P_{n+1} \\ & - [f_nI_n + \alpha_n + \alpha_{n+1}\gamma_n]P_n. \end{aligned} \quad (14.4.1)$$

Here P_n ($n = 0, 1, \dots$, where P_n with negative n is 0 by definition) denotes the probability for the molecular system to be in the n th RES M_n with free energy $F_n = \text{length}(M_n Q_n)$ at the configuration Q_n of the C.C. model of Fig. 14.2. The term $f_n I_n$ denotes the rate of an optical transition with oscillator strength f_n from this RES under solar radiation with intensity I_n at that transition energy (the definition of I and its value on the earth is given in the Appendix of Ref. [3]). The terms α_n and $\alpha_{n+1}\gamma_n$ denote the nonradiative transition rates for the de-excitation (to the state $n - 1$) from the same RES.

The stationary solution of (14.4.1) is readily given by

$$P_m = P_0 \prod_{n=1}^m [\gamma_{n-1} + f_{n-1}I_{n-1}/\alpha_n]. \quad (14.4.2)$$

Under the assumption of detailed balance (which does not necessarily hold except near thermal equilibrium), and ignoring the dependence on the concentrations of the reactants, we have the relation $\gamma_{n-1} = \exp[-(F_n - F_{n-1})/k_B T]$ with ambient temperature T . Putting $(F_n - F_{n-1}) \sim 1$ eV, we find γ_{n-1} as small as 10^{-20} , which means that the populations of the excited states in the dark ($I_n = 0$) are negligible, as it should. However, the solar radiation, although weak, gives rise to a significant effect. Putting the known intensity of solar radiation at the equator,⁶ averaged through day and night and throughout a year and over the energy range mentioned before, into the golden rule for the optical transition, one finds $I_n \sim 0.65 \text{ s}^{-1}$. The nonradiative de-excitation rate, given by the Eyring formula, $\alpha_n = \nu \exp(-\varepsilon_a/k_B T)$, can take a wide range of values depending upon the situation: ν , the frequency factor, can take values from 10^{11} s^{-1} to 10^3 s^{-1} while the exponent also is very sensitive to the activation energy ε_a which itself is typically a fraction of eV, being more than ten times greater than $k_B T$. In any case, one finds that the second term in $[\dots]$ of (14.4.2) is still much greater than the first, indicating that the solar radiation has a tremendous effect of populating the excited states, compared to the

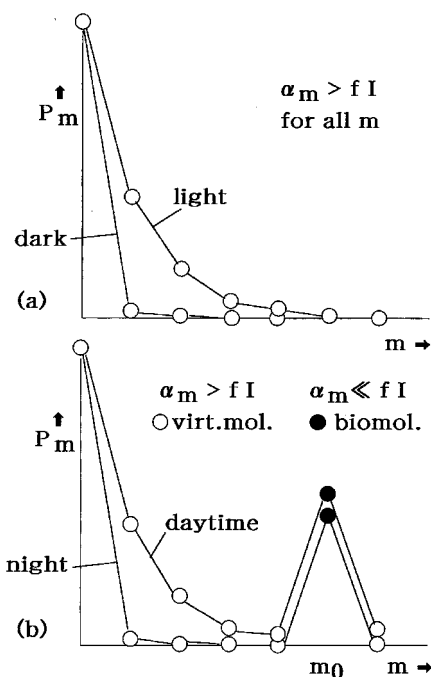


Fig. 14.3(a, b) Stationary solutions of the kinetic equations (14.4.1) for the populations of successive excited states in typical situations; (b) represents schematically the quasi-stationary states reached after a long time has elapsed since the primordial age.

dark case, as shown schematically in Fig. 14.3(a). For this reason, we hereafter neglect the first term in $[\dots]$ of eq. (14.4.2). Yet, the second term is much smaller than unity

$$f_{m-1}I_{m-1}/\alpha_m \ll 1, \quad (14.4.3)$$

so long as one puts the oscillator strength $f \sim 0.1$ and $\alpha > 1 \text{ s}^{-1}$, so that the population P_m is a rapidly decreasing function of m . The inequality (14.4.3) justifies our confinement of the kinetic equations to a single channel via the most long-lived intermediate states, since the survival chance of less long-lived excited states can be neglected.

On the other hand, we know that some of the simple biomolecules such as carbohydrates, amino acids and lipids have lifetimes α^{-1} much larger than seconds and possibly even than a day. If one assumes for simplicity that the inequality (14.4.3) is violated for $m \geq m_0$, P_m turns out to be an increasing function of m , and one cannot normalize the stationary population (14.4.2). This means that once m_0 is reached the molecule would grow indefinitely. In view of the existence of biopolymers of any length with long lifetime which violate the inequality (14.4.3), the stationary

solution of eq. (14.4.1) does not exist, implying that evolution proceeds indefinitely as long as the environment of the earth does not change drastically, in accordance with our present recognition of the living world.

A realistic assumption for *not too large* m may be that (14.4.3) is first violated at $m = m_0$ but is valid again for $m > m_0$, which results in two typical curves in Fig. 14.3 (b). What it implies will be mentioned later. In view of the *severe selectivity* (mentioned before) of successive photochemical reactions limited to the optimum condition, it is also possible that a *great number of identical* long-lived molecules are formed through parallel series of the same chains of reactions, especially *in the same environment* (possibly under catalytic regulation of unknown entity). This *might be* the physical mechanism underlying what is called *self-replication* of biomolecules. It is hardly conceivable, from a physical point of view, that a molecule replicates itself without an energy supply from outside, nor that the molecule knows the prescription for that self-replication even with that energy supply.

A meaningful quantity in such a situation is the flow ϕ from $m_0 - 1$ into m_0 , which is given by $f_{m_0-1} I_{m_0-1} \times P_{m_0-1}$ according to eq. (14.3.1). Making use of (14.4.2) with $m = m_0$, one obtains

$$\phi / P_0 = \prod_0^{m_0-1} f_n I_n / \prod_0^{m_0-1} \alpha_n. \quad (14.4.4)$$

By multiplying both sides by time t and by the number of starting molecules needed to form one biomolecule, which is of the order of m_0 , and equating them to 10%, we find that the time t needed to produce a quantity of biomolecules equal to 10% of the number of starting molecules is 0.2 billion years if one takes $f \sim 0.1$, the geometric mean $\langle \alpha \rangle \sim 10 \text{ s}^{-1}$ and $m_0 = 9$. This is not an unacceptable value in view of the duration of chemical evolution generally accepted, although we have too much allowance for the assumed values of the parameters. (Considering that the frequency factor ν is small due to slow and large relaxation in aqueous solution as will be described later and that we are selecting RES with the longest lifetime, we assumed $\langle \alpha \rangle$ to be smaller than would usually be expected.)

The molecules $m < m_0$ play an *indispensable role as intermediate steps* towards biomolecules, in spite of their shorter lifetimes and hence of their relatively low populations even in daytime. They will be called *virtual molecules*. When night comes, their population diminishes to insignificance within their short lifetimes, while the biomolecule population remains high even at night due to their long lifetimes, as shown schematically in Fig. 14.3(b). It is conceivable that the virtual molecules have so far been overlooked for this reason. A possible example of a virtual molecule will be mentioned later on. It should be emphasized that a significant number of biomolecules are *inevitably formed during a long span of time*, so long as their lifetimes are much larger than seconds.

14.5 Origin of natural selection and specific structure of biomolecules

We have seen that only a small number (“only one” in the simplified argument in the foregoing section) of the channels selected out of a great number of possible ones in multistep photochemical processes can lead successfully to the realization of biomolecules, other channels of molecular growth dropping out on the way. The similarity of this feature of the theory of chemical evolution to natural selection in biological evolution, and its dissimilarity to any events in nonliving systems which are not far from thermal equilibrium, leads us to support the postulate that natural selection itself started in the very beginning of chemical evolution. The absence of a clear-cut distinction between chemical and biological evolution lends further support to this viewpoint.

“Selected channels” of photochemical processes imply the “selected (and hence, specific) structures” of (virtual molecules and) biomolecules. Note that different optical transitions and different environments lead to different directions of relaxation in the multi-dimensional C.C. space, and hence generally to different RESs, namely, to different structures of newly-formed molecules in each step. The principle of selection is the longest-lived RESs or, more exactly speaking, the maximum probability of survival for the next photoexcitation. This principle of “maximal path probability” for the *evolving system* (far from equilibrium) is completely different in character from the principle of “minimal free energy” (equivalent to “maximal configurational probability”) governing systems in *thermal equilibrium*. In the latter case, the assemblage of (a few kinds of) a great number of identical atoms takes the form, at low temperatures, of a spatially (three-dimensionally) periodic array – the crystal lattice, as is realized by crystal growth from a melt or solution under a near-equilibrium situation – the better the equilibrium, the better the crystal! Conversely, in the former, the structures realized are very peculiarly shaped, some irregular chain-like array of atoms. Quasi-one-dimensional structures of biomolecules may be telling the history of their growth or evolution, while there is no trace of “time” recorded in a crystal, except for example in the screw dislocation – a one-dimensional structure reflecting the history of the crystal growth.

The fact that only twenty naturally-occurring species of amino acids are found in biomolecules among the many more species which can be synthesized in the laboratory may also be related to the different principles governing their formation as mentioned above. All species of amino-acid molecules in solution can be considered to be *metastable* states with a finite lifetime, corresponding to the *local* (instead of the absolute!) minima of the adiabatic potential surface on the multi-dimensional C.C. space of the assembly of constituent atoms, while very few of them, under solar radiation, have been realized through the selected channels from the prebiotic soup to be established as *real* biomolecules.

An indisputable argument based on a simplified model calculation shows how miraculous it is for a biomolecule with a complicated structure to be formed spontaneously from simple nonbiotic molecules.⁷ As for the mechanism for this miracle to be realized, D. Voet and J. G. Voet state it as follows,⁸ “How then did life arise? The answer, most probably, is that it was guided according to the survival of the fittest as it applies to the molecular level.” The scenario, presented here, of multi-channel photochemical reactions severely selected is a microscopic model which substantiates their speculation, from natural selection at the molecular level to the specific structures of biomolecules ascribable to it.

While the high chemical energy stored within molecules is an absolutely necessary condition for living activities, it is by no means a sufficient condition for that. However, it is physically and topologically obvious that long-chain biopolymers – a necessary consequence of the above scenario – have a great number of metastable configurations within a small interval of chemical energy and that this multistability could cause a variety of readily changeable functionalities of the biopolymers as is well known with the folding of protein molecules. It would be a long way, but not desperately long, from this multistability to the living activities. However, it should never be forgotten that all living activities, inclusive of self-replications and even the so-called free-will behaviors of animals, are driven by the energy of solar radiation indirectly as is obvious from the energy conservation law.

14.6 The roles of water in photochemical reactions

The vital importance of water for living organisms is obvious from the fact that more than half the weight of the human body consists of water. Water has been existent here on the surface of the earth ever since its formation and hence in the prebiotic era as well: there is no geological record that the *entire* water on the earth was ever frozen. The latter fact is attributed to the specific gravity of water which, as a function of temperature inclusive of its ice phase, takes its maximum at 4°C (instead of the monotonic decrease with temperature in normal substances). The bottom of the sea has always been kept at this temperature, even in the coldest era, with floating ice on the surface of the sea. This anomalous temperature dependence, especially the anomalous increase of specific gravity at melting, is due to the hydrogen bond between adjacent water molecules which constitutes the main part of the cohesive force in solid and liquid water. The hydrogen bond has a *particular orientation* with respect to the bent H₂O molecule, and an ice crystal (I_h) consists of an ordered network of oriented hydrogen bonds which is a rather spacious structure (opposite to the close-packed structure characteristic of van der Waals crystals). The order in the network is gradually destroyed as temperature rises, contributing a *finite decrease* in specific volume at the melting point and further *continuous decrease* up to 4°C.

Water is characterized also by its large specific heat (*per gram*) due to the small mass M of an H_2O molecule (note that *one atom* in condensed matter demands $3k_{\text{B}}$ specific heat at high enough temperature according to the Dulong–Petit law), which, together with the large total amount of water on the earth, contributes a great deal to keeping the surface temperature of the sea so moderate, somewhere between the freezing and boiling temperatures of water. The hydrogen bond, which is essentially due to the interaction between the *permanent* dipoles of the water molecule, is weaker than any of the ionic, covalent and metallic bonds. This is responsible for the low melting- and boiling-points of water compared to other crystals (except for van der Waals crystals which are formed by an even smaller interaction between *induced* dipoles of the molecules). As a result, under solar radiation, water evaporates readily which in turn contributes to keeping the temperature of the water left behind not too warm because of the heat of evaporation needed.

The weak hydrogen bond contributes to the large static dielectric constant of water, $\epsilon_s/\epsilon_0 \sim 80$, and hence to the large absorption coefficient in the infrared region as shown in Fig. 14.1(d). Most of solar radiation in the infrared region is absorbed near the surface of water, while that in the visible to near UV region can penetrate much deeper as is obvious empirically from the high transparency of clean water and is also shown in Fig. 14.1 with a vanishing absorption coefficient in this region. Photochemical reactions can take place at very deep locations in the clean sea and lakes, and at less deep locations in contaminated ponds, which contributes to equalizing the depth-integrated productivity.

The intense evaporation caused by the near-surface absorption of infrared light results in a column of heated vapor in the atmosphere, ascending to a height h where it counterbalances with the decreasing Boltzmann factor $\exp(-Mgh/k_{\text{B}}T)$ where g is the gravitational acceleration on the earth. (The H_2 molecules are so light and soared so high that most H atoms escaped the weak gravitational field of the earth in its early age except those in a compound form such as H_2O .) The water vapor is cooled due to the adiabatic expansion under decreasing pressure during ascension, becomes oversaturated and forms clouds, which will now drift horizontally towards lower pressure. Such convection of vapor brings about rain- and snow-falls elsewhere, resulting in river streams and waterfalls, contributing to the averaging and moderating of the climate (inclusive of moisture) between distant regions and different seasons. It is solar radiation that drives this circulating convection, which in turn provides environments most favorable to the living activities. A logically more proper statement may be that living organisms were created under, and evolved to adapt to, these environments. Nevertheless, the environments seem to be so wonderfully suited for the creation and evolution of living organisms that such would be seldom found elsewhere in the entire universe.

Let us now consider the more microscopic aspects of the roles of water in the creation of life. In this respect, the experimental study by Miller⁹ on the synthesis of biomolecules by an electric discharge in a gas mixture of simple abiotic molecules is very important. The high production of biomolecules in a relatively short time indicates that the molecular ions produced and accelerated by an electrostatic potential drop of several hundred volts which promoted their mutual collision played an essential role in driving the chemical reactions.

The photon energies of solar radiation are, at most, several electron volts, which are not enough to directly electrolyze water whose fundamental absorption band is above 8.3 eV (see Fig. 14.1(d)). However, some of the prebiotic molecules or ions which were presumably dissolved in the premordial sea or ponds may have had absorption bands within the spectral range of solar radiation. They could then not only play the role of a catalyst in electrolyzing water but could also form a chemical bond with solvent ions as will be speculated below.

Because of the extremely high dielectric constant, $\epsilon_S/\epsilon_0 \sim 80$, of water as a solvent, the screened Coulomb interaction, $\pm e^2/4\pi\epsilon_S r$, between two electronic charges at distance r is as small as 0.027 eV at $r \sim 5 \text{ \AA}$, being of the same order as the thermal energy $k_B T$ on the earth. This is a favorable situation *both in promoting and moderating* the chemical reactions between ionic and/or dipolar molecules which are of vital importance for living activities. However, the screening between two molecular ions or dipoles which are *adjacent* to each other in aqueous solution is much less effective due to the breakdown of the continuum model at this microscopic distance, the effective screening constant being somewhere between $1/\epsilon_0$ and $1/\epsilon_S$. For this reason, a molecular ion with small radius is combined with an adjacent water molecule with binding energy (point-dipole interaction) several to ten times greater than $k_B T$, forming, so to speak, a molecular cluster – the *hydration* effect. For example, a proton H^+ in water exists only in the form of the oxonium ion: H_3O^+ . It is to be noted in passing that the energy of the hydrogen bond in water consisting of dipolar (or almost point-dipole) interactions is 0.26 eV, as can be estimated from the vaporization energy of water.

Let us now remind ourselves of the fact that sodium chloride is automatically hydrolyzed into ions as it is dissolved in water. This takes place because the hydration energies of the anion and cation exceed the binding energy of the sodium-chloride crystal per pair of ions. Returning to our problem, the photoexcited electron (in LUMO – lowest unoccupied molecular orbital) of any solute molecule M has a much smaller binding energy than in the ground state (HOMO – highest occupied molecular orbit), and *might* be ionized to form a hydrated electron leaving the hydrated M^+ ion behind provided that the sum of their hydration energies exceeds the binding energy of the excited electron (see Fig. 14.4). A similar process, photo-induced self-decomposition of an electron and a hole,

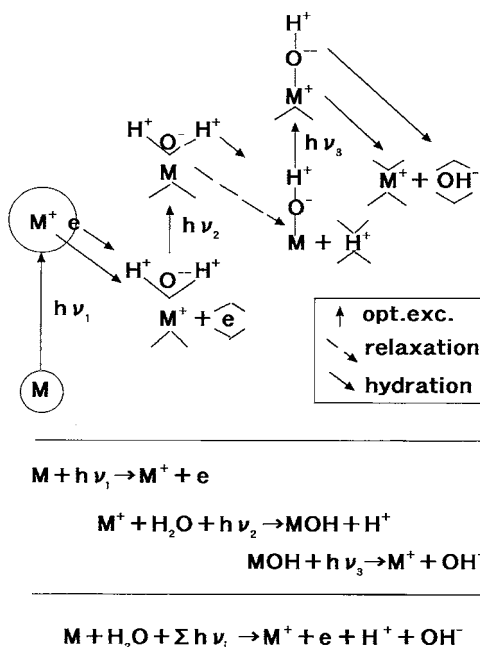


Fig. 14.4 A conceivable example of successive photochemical reactions (*upper*) which substantiates the unspecified scheme of Fig. 14.2 with the chemical reaction formulas shown for individual (*middle*) and entire (*lower*) steps. Hydration constitutes the main part of the relaxation here. Patterns like \wedge and \vee denote water molecules participating in the hydration, two of which are shown together with their atomic constituents: H^+ and O^- .

is known to arise even in condensed matter as was mentioned in Section 13.2 (see Fig. 13.3(c) and (d+d')). Because the distance from M^+ to an adjacent water molecule is contracted due to the hydration effect (this is a simplest example of the virtual molecule introduced in Section 14.4), it is conceivable that the electron on the water molecule is readily excited into the LUMO of M^+ (electron transfer) by absorbing the second photon, forming neutral M on one hand and breaking one of the O-H bonds of the water on the other. The corresponding H^+ would then be hydrated, leaving the compound MOH behind (one can expect a weak bond between M and OH since the new LUMO of M^+ extends to the O^- part of OH , O^- being still electronegative). The third photon would then excite the electron in the LUMO of M^+ back to OH , forming OH^- which would then be hydrated. The final products after the three-photon absorption are M^+ , H^+ , OH^- and an electron, all being in the hydrated state. If the electron returns to M^+ , which takes place without energy supply, the final result of the three-photon absorption is the decomposition of a water molecule into H^+ and OH^- , with M as a photocatalyst. The H^+ and OH^- ions will play important roles in further photochemical reactions.

The metallic ions such as Fe^{++} and Mn^{++} which were, like Na^+ , dissolved in the prebiotic soup due to natural hydration, are expected to play the same role as M^+ in the above series of reactions (so that one can start from the second step), and to contribute as catalyst to the multiphoton decomposition of water. It is interesting to note in this connection that the decomposition of water by four-photon absorption, with the Mn cluster as catalyst, has been observed in a more complicated photosystem II.¹⁰ As for the catalyzed photosynthesis of biomolecules, we refer to the study of Bard *et al.*¹¹ with titanium-oxide powder as the catalyst.

It is conceivable that a similar mechanism of photoliberation of H^+ and OH^- from solvent molecules is acting in the condensation polymerization of amino acids to protein and so on, which otherwise can take place only under limited conditions – a well-known difficulty in the experimental study of the chemical evolution of biomolecules to biopolymers. Recently, Matsuno and others¹² observed oligomerization of a peptide making use of a rapid cooling device similar to the hydrothermal vent. As for photopolymerization, Onoe and Kawai¹³ made use of powders of TiO_2 and other semiconductors as a photocatalyst.

A more concrete scenario of multistep photochemical reactions is left for future study.

14.7 Some remarks and further speculations

In the foregoing argument, we have tacitly assumed, as a first approximation, that solar radiation as the energy source, water as the medium of the multistep photochemical reactions and the atmosphere as the environment stayed unchanged ever since the origin of life. In fact, the solar radiation in the primordial time is supposed to have been somewhat less intense (except in the ultraviolet spectral region) but not so much as to affect our qualitative argument, and water has been essentially in the same situation as already mentioned.

In contrast, the atmospheric condition changed significantly: the partial pressure P of CO_2 in the primordial earth is supposed to have been much greater than at present while there was little O_2 before two billion years ago. Although the pressure-dependent part of the free energy (per molecule) of gas, $k_B T \ln P$,¹⁴ has no serious effect on our estimation of the chemical energy of biomolecules per atom ($\sim \text{eV}$) so long as the pressure change is within a thousand times, the scarcity of oxygen in an earlier era should be considered more carefully (this is the point we remarked on in Section 14.3). The living organisms in an early era were unable to perform such a wide range of oxygen-consuming living activities as at present although they were fortunately not exposed to dangerous attack by oxygen. Presumably, the chemical reactions and the way of acquiring energy in the origin and early evolution of living organisms were different from those in a later era, with much

lower activities. This is one of the points which make it difficult to extrapolate our arguments straightforwardly from the present back to the primordial time. It seems reasonable to confine our argument based on the estimation of chemical energy of biomolecules in Section 14.2 to the last two billion years, although prior to that we assume successive photochemical reactions with a *smaller* energy acquisition which evolved over time to the much more efficient photosynthesis of the present form. This evolution of functionality towards higher efficiency is nothing other than the result of natural selection which was built in at the very beginning of chemical evolution as described in Section 14.5. Darwinism is omnipresent in the living world.

According to recent studies,¹⁵ respiration was already started by the common ancestor of bacteria and archaea (before the appearance of photosynthesis, contrary to the earlier supposition) with the use of oxygen produced by abiotic photolysis of water, indicating the existence of dilute oxygen as early as more than three billion years ago. Then photosynthesis was started by some bacteria, initially without oxygen generation, and later on cyanobacteria or its ancestor started photosynthesis with oxygen generation.¹⁶ The oxygen thus generated in the water was initially used to oxidize Fe^{++} ions dissolved in water into Fe_2O_3 which was sedimented and piled as stratified deposits now found worldwide as the main source of iron. It was only after having oxidized the entire Fe ions in water (about 1.7 billion years ago) that the oxygen generated started to be stored in the atmosphere.

According to a quantitative estimation by Akiyama,¹⁷ the total carbon weights of living and dead organisms which amount to 4.57×10^{17} g and 3.7×10^{18} g, respectively, are much smaller than that of *organic* carbons such as coal and petroleum piled in the strata which amounts to 1.3×10^{22} g. Only 3.4% of the oxygen generated from inorganic material by photosynthesis exists as atmospheric oxygen, the remainder has been used to oxidize Fe ions in the sea as mentioned above. Of the present atmospheric oxygen, only a small part (0.1%) results from the photosynthesis of living (and already dead) organisms, the remainder coming from the photosynthesis of coal and petroleum. This is consistent with our proposed scenario of multistep photochemical reactions with severe competition for energy storage. Among the competitors, those products which piled as sediments at the *bottom* of the sea (the high transparency of water in the wide spectral region as seen in Fig. 14.1 may have helped this), being unable to react with the atmospheric oxygen for further living activities in spite of their success in storing energy, constitute the greatest part of the petroleum available at present. Here are also the (nonliving) organisms which have stored a great amount of solar energy in the form of chemical energy. In view of the enormous amount of petroleum available on the earth on the one hand and the severe competition among different channels on the other, the proposed scenario of identifying the petroleum with the “*never-born brothers*”

of living organisms seems more natural than the prevailing scenario identifying it with *dead* once-living organisms.

While natural selection in chemical evolution under solar radiation is a miniature version of that in biological evolution according to the retrospective analogy presented in Section 14.5, we suggest that the former developed into the latter according to the law of causality. If we are allowed to extend this sort of reasoning, we can even infer that the molecular processes taking place in chemical evolution (which can be well described by the physical principles) are the precursors or the embryos of the living activities appearing later on in biological evolution. This is borne out by a few examples presented in the following.

In the first place, we imagine that the multistep photochemical reactions proposed above as the origin of life evolved later to the much more efficient photosynthesis exhibited by cyanobacteria and finally to the present form of photosynthesis by plants. In proceeding from the second to the third stage, it seems, according to the symbiosis theory,¹⁸ that the photosynthesizing cyanobacteria evolved into another type of bacteria, endowing the photosynthetic function to the latter as the chloroplast of the eucaryotic cell which later evolved into plants – autotrophes. Another type of eucaryotic cell, formed by symbiosis of bacteria which had lost the function of photosynthesis such as mitochondrion, evolved into animals – heterotrophes. The former is directly driven by the solar radiation, while the latter, being fed by the former, is driven indirectly.

As a second example, the annealing of a photostructural change by thermal treatment or infrared absorption as seen in condensed matter (see channels 1 and 2 in Fig. 13.7) as well as a photochemical reaction (see channel α_1 and M_1RQ_0 in the left part of Fig. 14.2) is imagined to have evolved to the well-known biological fact that the fatal mutation in a gene caused by ultraviolet irradiation can sometimes be healed by thermal treatment with visible (or infrared) irradiation.¹⁹

For the third example, the fact that the human visible region coincides with the spectral region of intense solar radiation (see Fig. 14.1(c)) implies that visual pigments evolved under solar radiation so as to be better suited to survival in accordance with natural selection, which originated in the multistep photoreaction under solar radiation in its molecular stage.

Thus we expect that physical and chemical studies of the origin of life on the molecular level make a contribution to the final elucidation of the living activities in the biological stage which would not be obtained by retrospective studies alone.

In conclusion, we want to compare the photochemical evolution described in Sections 14.3, 14.4 with the thermodynamic Carnot cycle. Common to them is that the working material is driven between high- and low-temperature heat reservoirs (between the solar radiation with 5900 K and the earth environment with 290 K in the former). While the former is a *microscopic stochastic* process in the *open*

material system, the latter is a *macroscopic and deterministic* process with the *closed* working material. While the *chemical energy* is *stored within the matter* at each step of the photochemical reactions in the former, the *mechanical work* is done *upon the external system* in each Carnot cycle with *no storage of energy within the matter*. As a result, *the matter evolves chemically* in the former, performing living activities with the use of its stored chemical energy, while it does not evolve in the latter. However, the chance of *successful* photochemical change is *extremely small* in the former, resulting in the law of natural selection on the one hand and making the time-span of biological evolution very long on the other. The latter seems to be of vital importance for the appearance of highly intellectual animals.

In this sense, the living world is the most wonderful gift of the Heaven to the Earth.

Appendix 1

Derivation of exciton Hamiltonian from many-body Hamiltonian

In order to derive the matrix representation of the one-electron excitation energy, eq. (8.3.3) of the text, it is more convenient to make use of the second quantization formalism (see Chapter XIII of Ref. [2] in Chapter 2) than to resort to the algebra of Slater determinants as was done in the original paper by Wannier.¹ Denote by $\Psi(\mathbf{r})$ and $\Psi^\dagger(\mathbf{r})$ the operators for annihilation and creation, respectively, of an electron at position \mathbf{r} (c-number) which obey the anti-commutation rules for Fermi particles:

$$\begin{aligned} [\Psi(\mathbf{r}), \Psi^\dagger(\mathbf{r}')]_+ &\equiv \Psi(\mathbf{r})\Psi^\dagger(\mathbf{r}') + \Psi^\dagger(\mathbf{r}')\Psi(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}'), \\ [\Psi(\mathbf{r}), \Psi(\mathbf{r}')]_+ &= [\Psi^\dagger(\mathbf{r}), \Psi^\dagger(\mathbf{r}')]_+ = 0. \end{aligned} \quad (\text{A.1})$$

One can then write the Hamiltonian for a many-electron system as

$$\begin{aligned} \mathcal{H} = \mathcal{H}^{(1)} + \mathcal{H}^{(2)} &= \int d\mathbf{r} \Psi^\dagger(\mathbf{r}) h(\mathbf{p}, \mathbf{r}) \Psi(\mathbf{r}) \\ &+ (1/2) \iint d\mathbf{r} d\mathbf{r}' \Psi^\dagger(\mathbf{r}) \Psi^\dagger(\mathbf{r}') v(\mathbf{r} - \mathbf{r}') \Psi(\mathbf{r}') \Psi(\mathbf{r}) \end{aligned} \quad (\text{A.2})$$

with the use of the one-electron Hamiltonian $h(\mathbf{p}, \mathbf{r}) = -(\hbar^2/2m)\nabla^2 +$ (periodic potential due to all nuclei) and the interelectron Coulomb potential $v(\mathbf{r} - \mathbf{r}') = e^2/4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|$.

The operator $\Psi(\mathbf{r})$, as a function of the variable \mathbf{r} , can be expanded in terms of complete sets of orthonormal functions, firstly with the Bloch orbitals $\psi_{\nu\mathbf{k}}(\mathbf{r})$ and secondly with the Wannier orbitals $\phi_\nu(\mathbf{r} - \mathbf{R}_n)$ (the notation used only in this Appendix, instead of $a_\nu(\mathbf{r} - \mathbf{R}_n)$ used in Chapters 7 and 8, in order to distinguish it from the annihilation operator a defined below), as follows:

$$\Psi(\mathbf{r}) = \sum_{\nu\mathbf{k}} a_{\nu\mathbf{k}} \psi_{\nu\mathbf{k}}(\mathbf{r}) = \sum_{\nu n} a_{\nu n} \phi_\nu(\mathbf{r} - \mathbf{R}_n). \quad (\text{A.3})$$

where $a_{\nu\mathbf{k}}$ and $a_{\nu n}$ are annihilation operators of an electron in the Bloch state $(\nu\mathbf{k})$ and the Wannier state (νn) , respectively, and are related by

$$a_{\nu\mathbf{k}} = N^{-1/2} \sum_n \exp(-i\mathbf{k} \cdot \mathbf{R}_n) a_{\nu n} \quad (\text{A.4})$$

as can be confirmed with the use of (7.1.4). They satisfy the anti-commutation relations

$$[a_{\nu\mathbf{k}}, a_{\mu\mathbf{k}'}^\dagger]_+ = \delta_{\nu\mu} \delta_{\mathbf{k}\mathbf{k}'}, \quad [a_{\nu\mathbf{k}}, a_{\mu\mathbf{k}'}]_+ = [a_{\nu\mathbf{k}}^\dagger, a_{\mu\mathbf{k}'}^\dagger]_+ = 0 \quad (\text{A.5})$$

which follow from (A.1).

We will hereafter use band indices μ for the filled band, ν for the empty band, and λ for any band without such specification. The ground state of this system within the one-electron approximation is given by

$$|g\rangle = \prod_{\mu, \mathbf{k}} a_{\mu, \mathbf{k}}^\dagger |0\rangle, \quad (\text{A.6})$$

where $|0\rangle$ denotes the vacuum state and μ' runs over all the filled bands. Let us then consider the excited state $a_{\nu, \mathbf{k}}^\dagger a_{\mu, \mathbf{k}'} |g\rangle$ in which an electron has been raised from the (μ, \mathbf{k}') state to the (ν, \mathbf{k}) state. The total wave number \mathbf{K} of this state is given by $\mathbf{k} - \mathbf{k}'$. Since \mathbf{K} is a quantum number of the translationally symmetric Hamiltonian (A.2), an eigenstate $|e\rangle$ within the subspace of one-electron excitation can be written as a linear combination of the above-mentioned excited states with all possible sets of $(\mathbf{k}, \mathbf{k}')$ with a fixed value of $\mathbf{k} - \mathbf{k}' = \mathbf{K}$. We confine ourselves to a single set of (ν, μ) for Section 8.3, leaving more general treatments with linear combination on (ν, μ) for Section 8.5. Then we have

$$|e\rangle = B^\dagger |g\rangle \equiv \sum_{\mathbf{k}} f(\mathbf{k}) a_{\nu, \mathbf{k}}^\dagger a_{\mu, \mathbf{k}-\mathbf{K}} |g\rangle \quad (\text{A.7})$$

$$= N^{-1/2} \sum_{m\ell} \exp(i\mathbf{K} \cdot \mathbf{R}_m) F(\mathbf{R}_\ell) a_{\nu m+\ell}^\dagger a_{\mu m} |g\rangle. \quad (\text{A.7a})$$

The coefficients $f(\mathbf{k})$ and $F(\mathbf{R}_\ell)$, which are related through

$$f(\mathbf{k}) = N^{-1/2} \sum_{\ell} \exp(-i\mathbf{k} \cdot \mathbf{R}_\ell) F(\mathbf{R}_\ell), \quad (\text{A.8})$$

will be normalized as

$$\sum_{\mathbf{k}} |f(\mathbf{k})|^2 = \sum_{\ell} |F(\mathbf{R}_\ell)|^2 = 1. \quad (\text{A.9})$$

Equation (A.7a) represents the state in which an electron in the ν band and a hole in the μ band make relative motion with wave function $F(\mathbf{R}_\ell)$ and translational motion with wave vector \mathbf{K} – a reduction of the N -body problem to a two-body problem. What is then the Schrödinger equation for the wave function $F(\mathbf{R}_\ell)$ or $f(\mathbf{k})$ of the relative motion?

For a given Hamiltonian \mathcal{H} and a known ground state $|g\rangle$ of a many-body system, the following method for finding excited states $|e\rangle$ and excitation energy $E = E_e - E_g$ is in order: putting $|e\rangle = B^\dagger |g\rangle$ with an appropriate operator B^\dagger , one can rewrite the eigenequation for $|e\rangle$ as

$$0 = (\mathcal{H} - E_e)|e\rangle = (\mathcal{H}B^\dagger - E_e B^\dagger)|g\rangle = ([\mathcal{H}, B^\dagger]_- + B^\dagger \mathcal{H} - E_e B^\dagger)|g\rangle.$$

With the use of $\mathcal{H}|g\rangle = E_g|g\rangle$, the equation for B^\dagger is immediately obtained:

$$[\mathcal{H}, B^\dagger]_- |g\rangle = E B^\dagger |g\rangle. \quad (\text{A.10})$$

Multiplying $\langle g|a_{\mu, \mathbf{k}'}^\dagger a_{\nu, \mathbf{k}}$ by that equation obtained by putting B^\dagger of (A.7) into (A.10), one gets the equation for $f(\mathbf{k})$

$$\sum_{\mathbf{k}} H_{\mathbf{k}'\mathbf{k}} f(\mathbf{k}) = E f(\mathbf{k}') \quad (\text{A.11})$$

with the effective hamiltonian H defined by

$$H_{\mathbf{k}'\mathbf{k}} \equiv \langle g|a_{\mu, \mathbf{k}'}^\dagger a_{\nu, \mathbf{k}} [\mathcal{H}, a_{\nu, \mathbf{k}}^\dagger a_{\mu, \mathbf{k}-\mathbf{K}}]_- |g\rangle. \quad (\text{A.12})$$

Let us now calculate (A.12). Making use of the commutation relations

$$\begin{aligned} [\Psi(\mathbf{r}), a_{\nu, \mathbf{k}}^\dagger a_{\mu, \mathbf{k}-\mathbf{K}}]_- &= \phi_{\nu, \mathbf{k}}(\mathbf{r}) a_{\mu, \mathbf{k}-\mathbf{K}} \\ [\Psi^\dagger(\mathbf{r}), a_{\nu, \mathbf{k}}^\dagger a_{\mu, \mathbf{k}-\mathbf{K}}]_- &= -\phi_{\mu, \mathbf{k}-\mathbf{K}}^*(\mathbf{r}) a_{\nu, \mathbf{k}}^\dagger \end{aligned} \quad (\text{A.13})$$

obtained from (A.3) and (A.5) we can immediately calculate that part of the commutator in (A.12), which originates from the one-body Hamiltonian $\mathcal{H}^{(1)}$, as follows:

$$[\mathcal{H}^{(1)}, a_{\nu k}^\dagger a_{\mu k-K}]_- = \sum_{\lambda} (h_{\lambda \nu k} a_{\lambda k}^\dagger a_{\mu k-K} - h_{\mu \lambda k-K} a_{\nu k}^\dagger a_{\lambda k-K}). \quad (\text{A.14})$$

In deriving (A.14), we have put the matrix elements of h in regard to the Bloch state as $h_{\lambda k', \nu k} = \delta_{kk'} h_{\lambda \nu k}$ since the translationally symmetric h should be diagonal in \mathbf{k} .

Considering the occupancy in the ground state $|g\rangle$, we see that only $\lambda = \nu$ and $\lambda = \mu$ respectively in the first and the second terms of (A.14) contribute to (A.12), with the result:

$$H_{kk}^{(1)} = \delta_{kk} (h_{\nu \nu k} - h_{\mu \mu k-K}). \quad (\text{A.15})$$

The contribution of the two-body Hamiltonian $\mathcal{H}^{(2)}$ to the commutator in (A.12) can be calculated, by repeated use of (A.13) (twice with Ψ and twice with Ψ^\dagger), as

$$\begin{aligned} [\mathcal{H}^{(2)}, a_{\nu k}^\dagger a_{\mu k-K}]_- &= \int \int d\mathbf{r} d\mathbf{r}' \Psi^\dagger(\mathbf{r}) \Psi^\dagger(\mathbf{r}') v \Psi(\mathbf{r}') \phi_{\nu k}(\mathbf{r}) a_{\mu k-K} \\ &\quad - \int \int d\mathbf{r} d\mathbf{r}' \Psi^\dagger(\mathbf{r}) \phi_{\mu k-K}(\mathbf{r}') a_{\nu k}^\dagger v \Psi(\mathbf{r}') \Psi(\mathbf{r}), \end{aligned} \quad (\text{A.16})$$

which, in turn, is to be put into the expression for the effective Hamiltonian

$$H_{kk}^{(2)} = \langle g | a_{\mu k-K}^\dagger a_{\nu k} [\mathcal{H}^{(2)}, a_{\nu k}^\dagger a_{\mu k-K}]_- | g \rangle. \quad (\text{A.17})$$

With the use of the expansion (A.3), the contribution of the first term of (A.16) to (A.17) can be written as

$$\begin{aligned} &\sum_{\lambda_1 k_1 \lambda_2 k_2 \lambda_3 k_3} v_{\lambda_1 k_1 \lambda_2 k_2 \lambda_3 k_3 \nu k} \\ &\quad \times \langle g | a_{\mu k-K}^\dagger a_{\nu k} a_{\lambda_1 k_1}^\dagger a_{\lambda_2 k_2}^\dagger a_{\lambda_3 k_3} a_{\mu k-K} | g \rangle \end{aligned} \quad (\text{A.18})$$

with

$$\begin{aligned} &v_{\lambda_1 k_1 \lambda_2 k_2 \lambda_3 k_3 \lambda_4 k_4} \\ &\equiv \int \int d\mathbf{r} d\mathbf{r}' \phi_{\lambda_1 k_1}^*(\mathbf{r}) \phi_{\lambda_2 k_2}^*(\mathbf{r}') v(\mathbf{r} - \mathbf{r}') \phi_{\lambda_3 k_3}(\mathbf{r}') \phi_{\lambda_4 k_4}(\mathbf{r}). \end{aligned} \quad (\text{A.19})$$

It is convenient to study the diagonal ($\mathbf{k}' = \mathbf{k}$) and non-diagonal ($\mathbf{k}' \neq \mathbf{k}$) matrix elements, separately.

With $\mathbf{k}' = \mathbf{k}$, the expectation value $\langle g | \dots | g \rangle$ in (A.18) is nonvanishing only in the following two cases [the value being +1 in (i) and -1 in (ii)]:

- (i) $(\lambda_1 k_1) = (\nu k)$, $(\lambda_2 k_2) = (\lambda_3 k_3) \equiv (\mu' k'')$
- (ii) $(\lambda_2 k_2) = (\nu k)$, $(\lambda_1 k_1) = (\lambda_3 k_3) \equiv (\mu' k'')$

where $(\mu' k'')$ runs over all the occupied (in the ground state) Bloch states except $(\mu, \mathbf{k} - \mathbf{K})$. The corresponding contribution from the second term of (A.16) can be calculated in the same way. One gets altogether:

$$\begin{aligned} H_{kk}^{(2)} &= \sum_{\mu' k'' (\neq \mu k-K)} [(v_{\nu k \mu' k''} \mu' k'' \nu k - v_{\mu' k'' \nu k \mu k-K} \mu k-K) \\ &\quad - (v_{\mu' k'' \mu k-K \mu k-K} \mu k-K - v_{\mu' k'' \mu k-K \mu k-K} \mu k-K)]. \end{aligned} \quad (\text{A.20})$$

When $\mathbf{k}' \neq \mathbf{k}$, only the following terms are nonvanishing in (A.18):

- (i) $(\lambda_1 k_1) = (\nu k')$, $(\lambda_2 k_2) = (\mu k - \mathbf{K})$, $(\lambda_3 k_3) \equiv (\mu' k' - \mathbf{K})$,
- (ii) $(\lambda_1 k_1) = (\mu k - \mathbf{K})$, $(\lambda_2 k_2) = (\nu k')$, $(\lambda_3 k_3) \equiv (\mu' k' - \mathbf{K})$.

Since the second term of (A.16) contributes nothing, one immediately obtains

$$H_{\mathbf{k}'\mathbf{k}}^{(2)} = -v_{\nu\mathbf{k}'\mu\mathbf{k}-\mathbf{K}\mu\mathbf{k}'-\mathbf{K}\nu\mathbf{k}} + v_{\mu\mathbf{k}-\mathbf{K}\nu\mathbf{k}'\mu\mathbf{k}'-\mathbf{K}\nu\mathbf{k}} \quad (\mathbf{k}' \neq \mathbf{k}). \quad (\text{A.21})$$

Let us now choose the one-electron wave function $\phi_{\nu\mathbf{k}}(\mathbf{r})$ to be the self-consistent solutions of the Hartree–Fock equations:

$$\begin{aligned} h_{\text{H.F.}}\phi_{\lambda\mathbf{k}}(\mathbf{r}) \equiv & h(\mathbf{p}, \mathbf{r})\phi_{\lambda\mathbf{k}}(\mathbf{r}) + \sum_{\mu\mathbf{k}''} \int d\mathbf{r}' v(\mathbf{r} - \mathbf{r}')\phi_{\mu\mathbf{k}''}^*(\mathbf{r}') \\ & \times [\phi_{\mu\mathbf{k}''}(\mathbf{r}')\phi_{\lambda\mathbf{k}}(\mathbf{r}) - \phi_{\mu\mathbf{k}''}(\mathbf{r})\phi_{\lambda\mathbf{k}}(\mathbf{r}')] = \varepsilon_{\lambda}(\mathbf{k})\phi_{\lambda\mathbf{k}}(\mathbf{r}). \end{aligned} \quad (\text{A.22})$$

The solutions of (A.22) provide the Bloch functions not only for filled bands ($\lambda = \mu'$) with the occupancy (A.6) but also for empty bands ($\lambda = \nu'$). Being eigenfunctions of the common Hermite operator $h_{\text{H.F.}}$, they altogether form a complete orthonormal set. With the use of (A.19, A.22), one can write the one-electron energy as

$$\varepsilon_{\lambda}(\mathbf{k}) = h_{\lambda\lambda\mathbf{k}} + \sum_{\mu\mathbf{k}''} (v_{\lambda\mathbf{k}\mu\mathbf{k}''\mu\mathbf{k}''\lambda\mathbf{k}} - v_{\mu\mathbf{k}''\lambda\mathbf{k}\mu\mathbf{k}''\lambda\mathbf{k}}). \quad (\text{A.23})$$

The results (A.15, A.20, A.21) can be combined, with the help of (A.23), to give the explicit form of the effective Hamiltonian $H = H^{(1)} + H^{(2)}$:

$$\begin{aligned} H_{\mathbf{k}'\mathbf{k}} = & \delta_{\mathbf{k}'\mathbf{k}}[\varepsilon_{\nu}(\mathbf{k}) - \varepsilon_{\mu}(\mathbf{k} - \mathbf{K})] \\ & - v_{\nu\mathbf{k}'\mu\mathbf{k}-\mathbf{K}\mu\mathbf{k}'-\mathbf{K}\nu\mathbf{k}} + v_{\mu\mathbf{k}-\mathbf{K}\nu\mathbf{k}'\mu\mathbf{k}'-\mathbf{K}\nu\mathbf{k}}. \end{aligned} \quad (\text{A.24})$$

It implies the motion of an electron in the ν band and a positive hole in the μ band with attractive Coulomb and repulsive exchange interactions, with the total wave vector being kept constant ($[\mathbf{k} - (\mathbf{k} - \mathbf{K})] = \mathbf{K}$). Note that the signs of the two kinds of interactions in this electron–hole system are reversed as compared to the two-electron system.

Equation (A.24) represents a particular case of eq. (8.3.3) of the text in which both the conduction and valence bands are non-degenerate ($\mu = \mu'$, $\nu = \nu'$). The generalization to the degenerate case of eq. (8.3.3) itself will be left for the reader.

Appendix 2

Transient spectra for second-order optical processes

We have derived in Sections 11.4 and 11.5 the *stationary* and *time-resolved* spectra, respectively, of the second-order optical process of a localized electron. There is another type of spectrum studied by Kayanuma:² the *transient* spectrum in which one *switches on*, at $t = 0$, the excitation light with angular frequency Ω_1 (= photon energy since we put $\hbar = 1$ in those sections) and observe the response at a subsequent time t (under *continual excitation* for $t > 0$), in contradistinction to a *pulse excitation* at $t = 0$ in the case of the time-resolved spectrum. The transient spectrum has the merit of a *coexistent* light scattering component and a time-integrated (from 0 to t) emission spectrum at finite t (as the result of a suitable compromise under the uncertainty principle), as shown in Fig. A2.1(a) and (b) after a half-cycle and one cycle ($T \equiv 2\pi/\overline{\omega}$), respectively, of the interaction mode (as indicated by the clock) for the case of a strongly coupled localized electron. Note the coexistence of an increasingly resolved phonon structure of the light scattering component (left) and the *time-integrated* emission band (middle and right) reflecting the to-and-fro motion of the relaxing interaction mode.

The transient spectrum tends to a stationary spectrum in the limit of $t \rightarrow \infty$ as it should. Through this limiting procedure one can single out more directly the *hot-luminescence* component as shown by the broken curve in Fig. A2.2, although the component is included in the expression (11.4.36) and Fig. 11.5 for the stationary spectrum in a less explicit way.

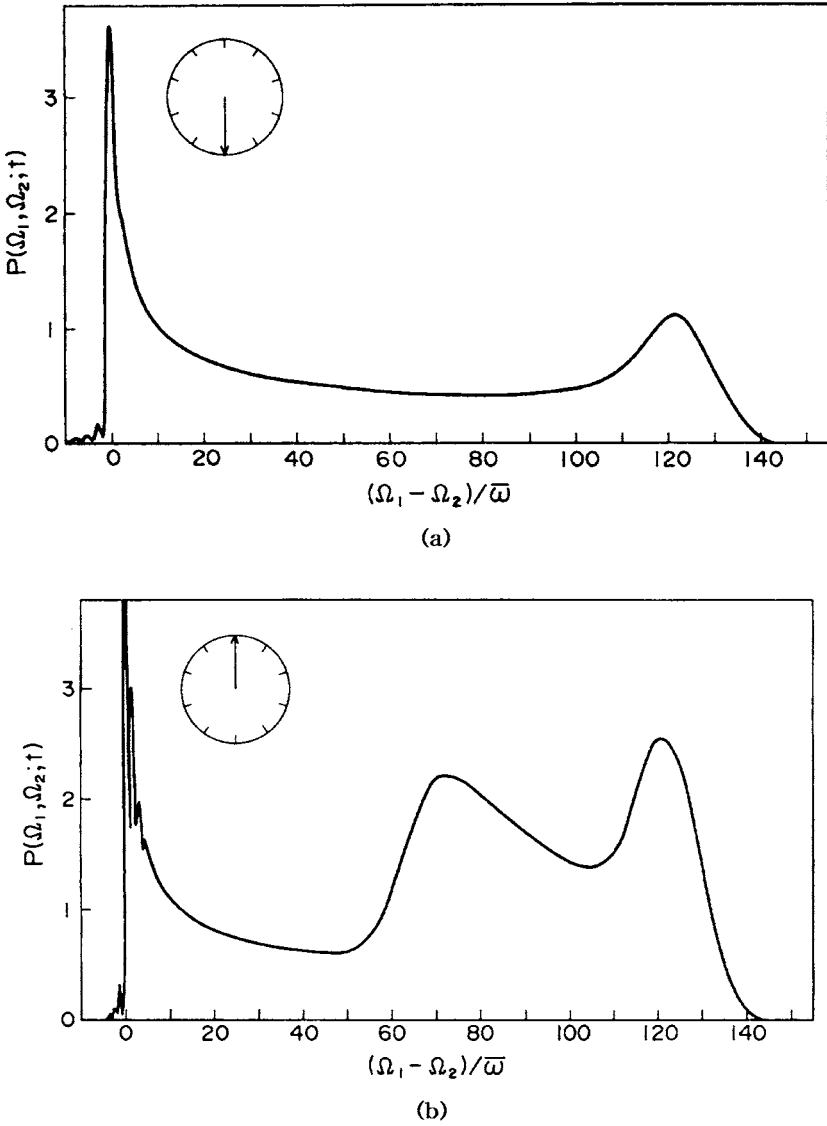


Fig. A2.1 The transient spectra of the resonant secondary radiation for a strongly coupled electron-phonon system, (a) at $t = 0.5T$ and (b) at $t = T$ after the switching on of the excitation light, where $T \equiv 2\pi/\bar{\omega}$ is one cycle of the interaction mode. The parameter values are $S = 40$, $\gamma/\bar{\omega} = 5 \times 10^{-5}$ and $\Omega_1 = E_a$. Due to Kayanuma.²

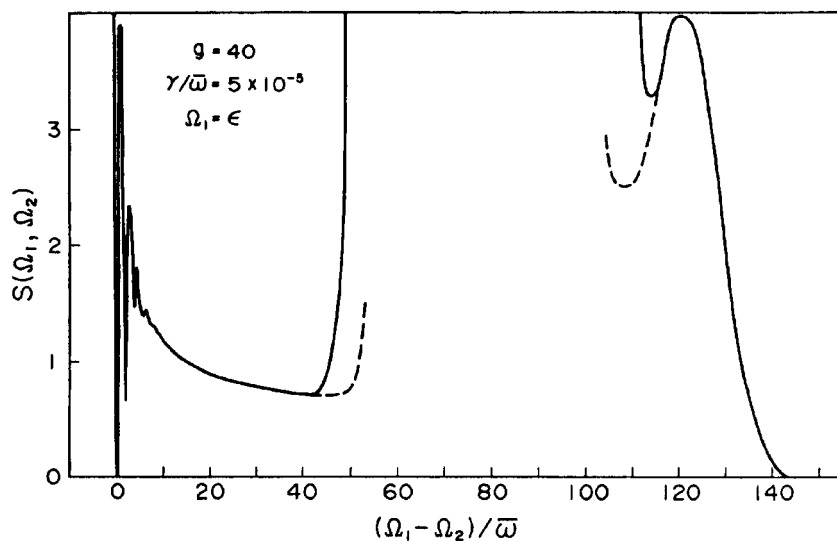


Fig. A2.2 The stationary spectrum of the resonant secondary radiation for a strongly coupled electron-phonon system. The hot-luminescence component is shown by broken lines. The parameter values are $S = 40$, $\gamma/\bar{\omega} = 5 \times 10^{-5}$ and $\Omega_1 = E_a$. Due to Kayanuma.²

References

Chapter 1

- 1 W. K. H. Panofsky and M. Phillips, *Classical Electricity and Magnetism*, 2nd edn (Addison-Wesley, Reading, Menlo Park, London, Sydney, Manila, 1962). A standard and useful textbook.
- 2 J. D. Jackson, *Classical Electrodynamics* (John Wiley and Sons, New York, London 1962). A textbook for an advanced theoretical course.
- 3 M. Planck, *Verh. dt. Phys. Ges.* **2**, 202; 237 (1900).

Chapter 2

- 1 P. A. M. Dirac, *Quantum Mechanics*, 4th edn (Oxford Univ. Press, London, 1958). A unique but clear presentation of general principles.
- 2 L. Schiff, *Quantum Mechanics*, 3rd edn (McGraw-Hill, New York, 1968). A standard textbook useful for various applications.
- 3 J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932). Hilbert space description of general principles, with logical analysis of the measurement problems in quantum mechanics.
- 4 Y. Toyozawa, *J. Phys. Soc. Jpn* **55**, 2572 (1986); K. Koide and Y. Toyozawa, *J. Phys. Soc. Jpn* **62**, 3385 (1993).

Chapter 3

- 1 A. Einstein, *Phys. Z.* **18**, 121 (1917).
- 2 R. Loudon, *The Quantum Theory of Light*, 2nd edn (Clarendon Press, Oxford, 1973). An extensive textbook on dynamics and statistics of the radiation field and its interaction with matter.
- 3 W. Heitler, *The Quantum Theory of Radiation*, 3rd edn (Clarendon Press, Oxford, 1954). A standard textbook on radiation-matter interaction.
- 4 See, for instance, Chapter VIII of Ref. [2.2].
- 5 H. Bethe, *Quantenmechanik der Ein- und Zweielektronenprobleme*, Chapter 3 of *Handbuch der Physik* **24/1** (Springer, Berlin, 1933).

Chapter 4

- 1 M. Born and J. R. Oppenheimer, *Ann. d. Phys.* **84**, 457 (1927).
- 2 R. Kubo and Y. Toyozawa, *Prog. Theor. Phys.* **13**, 160 (1955).

- 3 Y. Toyozawa and M. Inoue, *J. Phys. Soc. Jpn* **21**, 1663 (1966).
- 4 Y. Toyozawa, in *Dynamical Processes in Solid State Optics*, R. Kubo and H. Kamimura, eds. (Shokabo, Tokyo and W. A. Benjamin, Inc., New York, 1967), p. 90. See also, *J. Luminescence* **1**, **2**, 732 (1970).
- 5 Y. Toyozawa, A. Kotani and A. Sumi, *J. Phys. Soc. Jpn* **42**, 1495 (1977).
- 6 K. Huang, *Proc. Roy. Soc. A* **208**, 352 (1951).
- 7 M. Lax, *J. Chem. Phys.* **20**, 1753 (1952).
- 8 A. M. Stoneham, *Theory of Defects in Solids: Electronic Structure of Defects in Insulators and Semiconductors* (Clarendon Press, Oxford, 1985).
- 9 W. B. Fowler (ed.), *Physics of Color Centers* (Academic Press, New York, 1968).
- 10 H. A. Jahn and E. Teller, *Proc. Roy. Soc. A* **161**, 220 (1937).
- 11 T. Inui, Y. Tanabe, and Y. Onodera, *Group Theory and its Applications in Physics* (Springer, Berlin, Heidelberg, New York, Paris, Tokyo, Hong Kong, Barcelona, 1990).
- 12 H. C. Longuet-Higgins, U. Öpik, M. H. L. Pryce and R. A. Sack, *Proc. Roy. Soc. A* **244**, 1 (1958).
- 13 K. Cho, *J. Phys. Soc. Jpn* **25**, 1372 (1968).
- 14 P. H. Yuster and C. J. Delbecq, *J. Chem. Phys.* **21**, 892 (1953).
- 15 W.-U. Wagner, *Z. Phys.* **181**, 143 (1964) D. A. Patterson, *Phys. Rev.* **112**, 296 (1958).
- 16 S. G. Zazubovich, N. E. Lushchik and Ch. B. Lushchik, *Opt. Spectr.* **15**, 203 (1964).
- 17 A. Fukuda, K. Inohara and R. Onaka, *J. Phys. Soc. Jpn* **19**, 1964 (1964); A. Fukuda, *Science of Light* **13**, 37 (1964) (Thesis, Tokyo University of Education).
- 18 F. Seitz, *J. Chem. Phys.* **6**, 150 (1938).
- 19 S. Sugano, *J. Chem. Phys.* **36**, 122 (1962).
- 20 A. Fukuda, S. Makishima, T. Mabuchi and R. Onaka, *J. Phys. Chem. Solids*, **28**, 1763 (1967).
- 21 T. Shimada and M. Ishiguro, *Phys. Rev.* **187**, 1089 (1969).
- 22 S. Masunaga, N. Goto, A. Matsushima and A. Fukuda, *J. Phys. Soc. Jpn* **43**, 2013 (1977).
- 23 Y. Kamishina, V. S. Shibasankar and P. W. M. Jacobs, *J. Chem. Phys.* **76**, 4677 (1982); V. S. Shibasankar, Y. Kamishina and P. W. M. Jacobs, *ibid.*, 4681 (1982).
- 24 K. Ueda, Y. Shimizu, H. Chiba, M. Okunishi, K. Ohmori, Y. Sato, F. Shigemasa and N. Kosugi, *J. Electron Spectr.* **79**, 441 (1996).
- 25 S. Itoh, S. Tanaka and Y. Kayanuma, *Phys. Rev. A* **60**, 4488 (1999).
- 26 R. E. Dietz, D. G. Thomas and J. J. Hopfield, *Phys. Rev. Lett.* **8**, 391 (1962).
- 27 M. D. Sturge, *Phys. Rev.* **130**, 639 (1963).

Chapter 5

- 1 C. Kittel, *Introduction to Solid State Physics*, 7th edn (John Wiley & Sons, Inc., New York, Chichester, Brisbane, Toronto, Singapore, 1996). One of the standard textbooks.
- 2 M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Clarendon Press, Oxford, 1954). A standard textbook.
- 3 P. W. Anderson, *Basic Notion of Condensed Matter Physics* (Benjamin/Cummings, Menlo Park, London, Amsterdam, Don Mills, Sydney, Tokyo, 1984).
- 4 H. Fröhlich, *Theory of Dielectrics* (Clarendon Press, Oxford, 1949).
- 5 J. H. Van Vleck, *The Theory of Electric and Magnetic Susceptibilities* (Clarendon Press, Oxford, 1932), Chapter I. See also Chapters 2–3 of Ref. [1.1].
- 6 T. Kurosawa, *J. Phys. Soc. Jpn* **16**, 1299 (1961).

- 7 M. Cardona (ed.), *Light Scattering in Solids I* (Springer Verlag, Berlin, Heidelberg, New York, 1983).
- 8 W. Cochran and R. A. Cowley, Phonons in perfect crystals, in *Light and Matter Ia. Encyclopedia in Physics, Vol. XXV/2a*, S. Flügge, ed. (Springer, Berlin, Heidelberg, New York, 1967), p. 59.
- 9 M. Okazaki, M. Inoue, Y. Toyozawa, T. Inui and E. Hanamura. *J. Phys. Soc. Jpn* **22**, 1349 (1967).

Chapter 6

- 1 R. Kubo, *J. Phys. Soc. Jpn* **12**, 570 (1957).
- 2 H. Nakano, *Prog. Theor. Phys.* **15**, 77 (1956).
- 3 Y. Toyozawa, Polarization waves and dielectric dispersion, in *The Physics of Elementary Excitations*, S. Nakajima, Y. Toyozawa and R. Abe, eds. (Springer, Berlin, Heidelberg, New York, 1980), Chapter 2.
- 4 D. Pines (ed.), *The Many Body Problem* (Benjamin, New York, 1962).
- 5 M. Born, *Optik* (Springer, Berlin, 1933), Sections 83 and 84.

Chapter 7

- 1 R. E. Peierls, *Quantum Theory of Solids* (Clarendon Press, Oxford, 1964).
- 2 N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976).
- 3 W. Shockley, *Electrons and Holes in Semiconductors* (van Nostrand, New York, Toronto, London, 1950).
- 4 W. Kohn, Shallow impurity states in silicon and germanium, in *Solid State Physics*, Vol. 5, F. Seitz and D. Turnbull, eds. (Academic Press, New York, London, 1957).

Chapter 8

- 1 R. S. Knox, Theory of excitons, *Solid State Physics, Suppl.*, Vol. 5, F. Seitz and D. Turnbull, eds. (Academic Press, New York, 1953). A well-organized review article of early developments.
- 2 J. Frenkel, *Phys. Rev.* **37**, 17, 1276 (1931).
- 3 W. R. Heller and A. Marcus, *Phys. Rev.* **84**, 809 (1951).
- 4 A. S. Davydov, *Theory of Molecular Excitons* (Plenum Press, New York, 1971).
- 5 G. H. Wannier, *Phys. Rev.* **52**, 191 (1937).
- 6 R. J. Elliott, *Phys. Rev.* **108**, 1384 (1957).
- 7 T. Tomiki, *J. Phys. Soc. Jpn* **26**, 738 (1969).
- 8 S. Nikitine, J. B. Grun and M. J. Sieskind, *J. Phys. Chem. Solids* **17**, 292 (1961).
- 9 Y. Onodera and Y. Toyozawa, *J. Phys. Soc. Jpn* **22**, 833 (1967).
- 10 Y. Kato, C. I. Yu and T. Goto, *J. Phys. Soc. Jpn* **28**, 104 (1970).
- 11 T. Koda and D. W. Langer, *Phys. Rev. Lett.* **20**, 50 (1968).
- 12 O. Akimoto and H. Hasegawa, *Phys. Rev. Lett.* **20**, 916 (1968).
- 13 T. Koda, T. Murahashi, T. Mitani, S. Sakoda and Y. Onodera, *Phys. Rev.* **B5**, 705 (1972).
- 14 K. Cho, Internal structure of excitons, in *Excitons*, K. Cho, ed. (Springer, Berlin, Heidelberg, New York, 1979), Chapter 2.
- 15 M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Clarendon Press, Oxford, 1954).
- 16 J. J. Hopfield, *Phys. Rev.* **112**, 155 (1958).

- 17 Y. Onodera and Y. Toyozawa, *J. Phys. Soc. Jpn* **24**, 341 (1968).
- 18 D. W. Taylor, *Phys. Rev.* **156**, 1017 (1967).
- 19 P. Soven, *Phys. Rev.* **156**, 809 (1967).
- 20 B. Velicky, S. Kirkpatrick and H. Ehrenreich, *Phys. Rev.* **175**, 747 (1968).
- 21 E. I. Rashba and G. E. Gurgenishvili, *Sov. Phys.-Solid State* **4**, 759 (1962).
- 22 C. H. Henry and K. Nassau, *J. Luminescence* **1**, **2**, 299 (1970).
- 23 Y. Nakai, T. Murata and K. Nakamura, *J. Phys. Soc. Jpn* **18**, 481 (1963); K. Nakamura and Y. Nakai, *J. Phys. Soc. Jpn* **23**, 455 (1967).
- 24 T. Murata and Y. Nakai, *J. Phys. Soc. Jpn* **23**, 904 (1967).
- 25 H. Kanzaki, in *Excitonic Processes in Solids*, M. Ueta, H. Kanzaki, K. Kobayashi, Y. Toyozawa and E. Hanamura (Springer, Berlin, Heidelberg, New York, Tokyo, 1984), Chapter 6.
- 26 K. Takahei and K. Kobayashi, *J. Phys. Soc. Jpn* **43**, 891 (1977).
- 27 S. D. Mahanti, *Phys. Rev.* **B10**, 1384 (1974).
- 28 N. Nagasawa, N. Nakagawa and Y. Nakai, *J. Phys. Soc. Jpn* **24**, 1403 (1968).
- 29 V. L. Broude, E. I. Rashba and E. F. Sheka, *Spectroscopy of Molecular Excitons* (Springer, Berlin, Heidelberg, New York, Tokyo, 1985).

Chapter 9

- 1 L. D. Landau, *Phys. Z. Sowietunion*, **3**, 664 (1933).
- 2 S. I. Pekar, *Untersuchungen über die Elektronentheorie der Kristalle* (Akademie-Verlag, Berlin, 1954).
- 3 H. Fröhlich, *Adv. in Phys.* **3**, 325 (1954).
- 4 T. D. Lee and D. Pines, *Phys. Rev.* **88**, 960 (1952).
- 5 R. P. Feynman, *Phys. Rev.* **97**, 660 (1955).
- 6 C. G. Kuper and G. D. Whitfield (eds.), *Polarons and Excitons* (Oliver and Boyd, Edinburgh, 1963).
- 7 J. T. Devreese (ed.), *Polarons in Ionic Crystals and Polar Semiconductors* (North-Holland, Amsterdam, 1972).
- 8 Y. Toyozawa, *Dynamical Properties of Charge Carriers in Dielectrics – Generalization of the Polaron Theory*, in Ref. [7], p.1.
- 9 M. Gurari, *Phil. Mag.* **44**, 329 (1953).
- 10 S. V. Tiablikov, *Zh. Eksp. i. Teor. Fiz.* **22**, 325 (1952); *ibid.*, **23**, 381 (1952); *ibid.*, **26**, 545 (1954).
- 11 T. D. Lee, F. Low and D. Pines, *Phys. Rev.* **90**, 297 (1953).
- 12 R. P. Feynman, *Rev. Mod. Phys.* **20**, 367 (1948).
- 13 A. S. Mishchenko, N. V. Prokof'ev, A. Sakamoto and B. V. Svistunov, *Phys. Rev.* **B62**, 6317 (2000).
- 14 I. D. Feranchuk, S. I. Fisher and L. I. Komarov, *J. Phys. C* **18**, 5083 (1985).
- 15 Y. Toyozawa, *Prog. Theor. Phys.* **12**, 421 (1954).
- 16 W. B. Fowler, *Phys. Rev.* **151**, 657 (1966).
- 17 M. Inoue, C. K. Mahutte and S. Wang, *Phys. Rev.* **B2**, 539 (1970).
- 18 A. B. Kunz, *Phys. Rev.* **B2**, 5015 (1970); N. O. Lipari and A. B. Kunz, **B3**, 491(1971); A. B. Kunz and N. O. Lipari, **B4**, 1374, 4639 (1971); A. B. Kunz, **B6**, 606 (1972); A. B. Kunz, **B7**, 5369 (1973); A. B. Kunz and D. J. Mickish, **B8**, 779 (1973); A. B. Kunz, **B8**, 1690 (1973); D. J. Mickish, A. B. Kunz and T. C. Collins, **B9**, 4461 (1974); S. T. Pantelides, D. J. Mickish and A. B. Kunz, **B10**, 2602, 5203 (1974); A. B. Kunz and D. J. Mickish, **B11**, 1700 (1975); A. B. Kunz, **B12**, 5890 (1975);
- 19 S. Suhai, *Phys. Rev.* **B29**, 4570 (1984).

- 20 H. Haken, in *Polarons and Excitons*, C. G. Kuper and D. G. Whitfield, eds. (Oliver and Boyd, Edinburgh, 1963), p. 302.
- 21 T. Miyakawa, *J. Phys. Soc. Jpn* **17**, 1898 (1962).
- 22 J. C. Hermanson, *Phys. Rev.* **177**, 1234 (1969).
- 23 J. T. Devreese, A. B. Kunz and T. C. Collins, *Solid State Commun.* **11**, 673 (1972); see also, Ref. [7], p. 735.
- 24 B. Steeg, E. Gminder, M. Kirm, V. Kisand, S. Vielhauer and G. Zimmerer, *J. Electron Spectr.* **101–103**, 879 (1999).
- 25 S. Vielhauer, E. Gminder, M. Kirm, V. Kisand, E. Negodin, B. Steeg and G. Zimmerer, *Abstracts of the 4th Int. Conf. of Excitonic Processes in Condensed Matter (Osaka, 2000)*, p. 94.
- 26 W. Känzig, *Phys. Rev.* **99**, 1890 (1955); T. O. Woodruff and W. Känzig, *J. Phys. Chem. Solids* **5**, 268 (1958).
- 27 J. Bardeen and W. Shockley, *Phys. Rev.* **80**, 72 (1950).
- 28 Y. Toyozawa, *Prog. Theor. Phys.* **26**, 29 (1961). See also Y. Toyozawa, Self-trapping of an electron by the acoustic mode of lattice vibration, in Ref. [6], p. 211.
- 29 H. Sumi, *J. Phys. Soc. Jpn* **36**, 770 (1974).
- 30 Y. Toyozawa, in *Relaxation of Elementary Excitations*, R. Kubo and E. Hanamura, eds. (Springer, Berlin, Heidelberg, New York, 1980), p. 3.
- 31 R. Kubo, *J. Phys. Soc. Jpn* **3**, 254 (1948); **4**, 322 (1949).
- 32 E. I. Rashba, *Opt. Spectr.* **2**, 75, 88 (1957).
- 33 E. I. Rashba, Self-trapping of excitons, in *Excitons*, E. I. Rashba and M. D. Sturge, eds. (North Holland, Amsterdam 1982), p. 543.
- 34 D. Emin and T. Holstein, *Phys. Rev. Lett.* **36**, 323 (1976).
- 35 L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Inc., New York, Toronto, London, 1949).
- 36 A. Sumi and Y. Toyozawa, unpublished.
- 37 S. Higai and H. Sumi, *J. Phys. Soc. Jpn* **63**, 4489 (1994).
- 38 B. Pertzsch and U. Rössler, *Solid State Commun.* **37**, 931 (1981).
- 39 A. Sumi, *J. Phys. Soc. Jpn* **43**, 1286 (1977).
- 40 M. N. Kabler, Hole centers in halide lattices, in *Point Defects in Solids, Vol. 1*, J. H. Crawford, Jr, and L. M. Slifkin, eds., (Plenum, New York 1972), p. 327.
- 41 Ch. B. Lushchik, Free and self-trapped excitons in alkali halides: spectra and dynamics, in *Excitons*, E. I. Rashba and M. D. Sturge, eds. (North Holland, Amsterdam 1982), p. 543.
- 42 T. Hayashi, T. Ohata and S. Koshino, *J. Phys. Soc. Jpn* **42**, 1647 (1977).
- 43 H. Nishimura, H. Miyazaki, Y. Tanaka, K. Uchida and M. Tomura, *J. Phys. Soc. Jpn* **48**, 1829 (1979).
- 44 T. V. Khiem, A. Nouhailhat, *J. Phys. Soc. Jpn* **50**, 127 (1981).
- 45 A. Matsui, M. Iemura and H. Nishimura, *J. Luminescence* **24/25**, 4454 (1981).
- 46 H. Kanzaki, S. Sakuragi and K. Sakamoto, *Solid State Commun.* **9**, 999 (1971).
- 47 R. C. Hanson, *J. Phys. Chem.* **66**, 2376 (1962).
- 48 T. Kawai, K. Kobayashi, M. Kurita and Y. Makita, *J. Phys. Soc. Jpn* **30**, 1101 (1971).
- 49 Y. Toyozawa and A. Sumi, in *Proc. 12th Int. Conf. Phys. Semicond., Stuttgart (1974)*, M. H. Pilkuhn, ed. (Teubner, Stuttgart 1974), p. 179.
- 50 K. Takahei and K. Kobayashi, *J. Phys. Soc. Jpn* **43**, 891 (1977).
- 51 Y. Shinozuka and Y. Toyozawa, *J. Phys. Soc. Jpn* **46**, 505 (1979).
- 52 Y. Toyozawa, *Solid-State Electronics* **21**, 1313 (1978).
- 53 K. S. Song and R. T. Williams, *Self-Trapped Excitons* (Springer, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hongkong, Barcelona, Budapest, 1993).
- 54 I. Ya. Fugol', *Adv. Phys.* **27**, 1 (1978); **37**, 1 (1988).

- 55 H. Kanzaki, Excitons in condensed rare gases, in *Excitonic Processes in Solids* M. Ueta, H. Kanzaki, K. Kobayashi, Y. Toyozawa and E. Hanamura (Springer, Berlin, Heidelberg, New York, Tokyo, 1986), Chapter 5. It contains a number of references on condensed rare gases relevant to the descriptions in Section 9.6.
- 56 J. Poitrenaud and F. I. B. Williams, *Phys. Rev. Lett.* **29**, 1230 (1972).
- 57 W. B. Fowler and D. L. Dexter, *Phys. Rev.* **176**, 337 (1968).
- 58 K. R. Atkins, *Phys. Rev.* **116**, 1339 (1959).
- 59 S. V. Nistor, E. Goovaerts and D. Shoemaker, *Phys. Rev.* **B48**, 9575 (1993).
- 60 T. Hirota, T. Fujita and Y. Kazumata, *Jpn J. Appl. Phys.*, **32**, 4674 (1993).
- 61 T. Suemoto and H. Kanzaki, *J. Phys. Soc. Jpn* **50**, 3664 (1980).
- 62 C. H. Leung, G. Brunet and K. S. Song, *J. Phys. C* **18**, 4459 (1985). For more details of the background, see also Section 5.2.3 of Ref. [53].
- 63 K. Kan'no, K. Tanaka and T. Hayashi, *Reviews of Solid State Science*, **4**, 383 (1990) (World Scientific); T. Matsumoto, T. Kawata, A. Miyamoto and K. Kan'no, *J. Phys. Soc. Jpn* **61**, 4229 (1992); T. Matsumoto, M. Shirai and K. Kan'no, *ibid.* **64**, 987 (1995); K. Kan'no, T. Matsumoto and Y. Kayanuma, *Pure & Appl. Chem.* **69**, 1227 (1997).
- 64 M. Iwanaga, M. Watanabe and T. Hayashi, *Phys. Rev.* **B62**, 10766 (2000); Also, private communication by T. Hayashi.
- 65 M. Höhne and M. Stasiw, *phys. stat. solidi*, **28**, 247 (1968). See also Section 6.2 of the book in Ref. [55].

Chapter 10

- 1 A. I. Ansel'm, Iu. A. Firsov, *J. Exp. Theor. Phys.* **28**, 151 (1955); **30**, 719 (1956).
- 2 Y. Toyozawa, *Prog. Theor. Phys.* **20**, 53 (1958).
- 3 Y. Toyozawa and Y. Shinozuka, *J. Phys. Soc. Jpn* **48**, 472 (1980).
- 4 H. Haken, *Zeits. f. Phys.* **146**, 527 (1956); **147**, 323 (1957); also in *Polarons and Excitons*, C. G. Kuper and D. G. Whitfield, eds. (Oliver and Boyd, Edinburgh, 1963), p. 302.
- 5 M. Matsuura and H. Büttner, *Phys. Rev.* **B21**, 679 (1980).
- 6 J. Pollmann and H. Büttner, *Solid State Commun.* **17**, 1171 (1975); *Phys. Rev.* **B16**, 4480 (1977).
- 7 T. P. McLean, The absorption edge spectrum of semiconductors, in *Progress of Semiconductors*, Vol. 5, A. F. Gibson, F. A. Kröger and R. E. Burgess, eds. (Wiley, New York, 1961), p. 53.
- 8 Y. Toyozawa, *J. Phys. Chem. Solids* **25**, 59 (1964).
- 9 Y. Toyozawa, *J. Luminescence*, **1**, 2, 732 (1970).
- 10 F. C. Brown, T. Masumi and H. H. Tippins, *J. Phys. Chem. Solids* **22**, 101 (1961).
- 11 J. Nakahara, K. Kobayashi and A. Fujii, *J. Phys. Soc. Jpn* **37**, 1312 (1974).
- 12 P. W. Anderson, *J. Phys. Soc. Jpn* **9**, 316 (1954).
- 13 R. Kubo, *J. Phys. Soc. Jpn* **9**, 935 (1954).
- 14 Y. Toyozawa, Localization and delocalization of an exciton in the phonon field, in *Organic Molecular Aggregates*, P. Reineker, H. Haken and H. Wolf, eds. (Springer, Berlin, Heidelberg, New York, Tokyo, 1983), p. 90.
- 15 Y. Toyozawa, *Prog. Theor. Phys.* **27**, 89 (1962).
- 16 H. Sumi, *J. Phys. Soc. Jpn* **32**, 616 (1972).
- 17 T. Itoh, S. Watanabe and M. Ueta, *J. Phys. Soc. Jpn* **48**, 542 (1980).
- 18 T. Miyata, *J. Phys. Soc. Jpn* **31**, 529 (1971).
- 19 G. Baldini, A. Bosacchi and B. Bosacchi, *Phys. Rev. Lett.* **23**, 846 (1969).

- 20 Y. Toyozawa and J. Hermanson, *Phys. Rev. Lett.* **21**, 1637 (1968).
- 21 J. C. Hermanson, *Phys. Rev.* **B2**, 5043 (1970).
- 22 W. Y. Liang and A. D. Yoffe, *Phys. Rev. Lett.* **20**, 59 (1968).
- 23 W. C. Walker, D. N. Roessler and E. Loh, *Phys. Rev. Lett.* **20**, 847 (1968).
- 24 R. Z. Bachrach and F. C. Brown, *Phys. Rev. Lett.* **21**, 685 (1968); *Phys. Rev.* **B1**, 818 (1970).
- 25 S. Kurita and K. Kobayashi, *J. Phys. Soc. Jpn* **30**, 1645 (1971).
- 26 M. Kamada, M. Yoshikawa and R. Kato, *J. Phys. Soc. Jpn* **39**, 1004 (1975).
- 27 M. Kamada, T. Yoshimura and R. Kato, *J. Phys. Soc. Jpn* **42**, 1660 (1977).
- 28 G. Dolling, J. Sakurai and R. A. Cowley, *Proc. Int. Meeting on Ferroelectricity, Kyoto, 1969, J. Phys. Soc. Jpn* **28**, Suppl., p. 258 (1970).
- 29 T. Sakai, H. Kawaura and R. Kato, *J. Phys. Soc. Jpn* **56**, 1943 (1987).
- 30 M. Ashida, Y. Kawaguchi and R. Kato, *J. Phys. Soc. Jpn* **58**, 4620 (1989).
- 31 A. Kato, M. Ashida and R. Kato, *J. Phys. Soc. Jpn* **66**, 2886 (1997).
- 32 T. Holstein, S. K. Lyo and R. Orbach, *Laser Spectroscopy in Solids* (Springer, Berlin, 1981), p. 39.
- 33 H. Kawaura, Y. Kawaguchi and R. Kato, *J. Phys. Soc. Jpn* **57**, 3613 (1988).
- 34 F. Urbach, *Phys. Rev.* **92**, 1324 (1953).
- 35 M. V. Kurik, An extensive review article on the *Urbach Rule*, in *phys. status solidi (a)* **8**, pp. 9–45 (1971).
- 36 U. Haupt, *Z. Phys.* **157**, 232 (1959).
- 37 M. Schreiber and Y. Toyozawa, *J. Phys. Soc. Jpn* **51**, 1528, 1537, 1544 (1982); *ibid.* **52**, 318 (1983); *ibid.* **53**, 864 (1984).
- 38 D. L. Dexter, *Nuovo Cimento, Suppl.* **7**, 45 (1958).
- 39 Y. Toyozawa, *Prog. Theor. Phys.* **22**, 445 (1959).
- 40 H. Mahr, *Phys. Rev.* **125**, 1510 (1962).
- 41 Y. Toyozawa, *Tech. Rep. ISSP, Ser. A*, No. 119 (1964).
- 42 P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
- 43 H. Sumi and Y. Toyozawa, *J. Phys. Soc. Jpn* **31**, 342 (1971).
- 44 H. Sumi, *J. Phys. Soc. Jpn* **32**, 616 (1972).
- 45 J. D. Dow and D. Redfield, *Phys. Rev.* **B5**, 594 (1972).
- 46 E. Mohler and B. Thomas, *Phys. Rev. Lett.* **4**, 543 (1980).
- 47 Y. Toyozawa, *Proc. IV. Int. Conf. on Vacuum Ultraviolet Radiation Physics, Hamburg, 1974* (Pergamon, Vieweg, 1974), 317.
- 48 T. Tomiki, *J. Phys. Soc. Jpn* **22**, 463 (1967); T. Tomiki, T. Miyata and H. Tsukamoto, *J. Phys. Soc. Jpn* **35**, 495 (1973).
- 49 D. M. Burland, U. Konzelmann and R. M. Mcfarlane, *J. Chem. Phys.* **67**, 1962 (1977); D. M. Burland, D. E. Cooper, M. D. Fayer and C. R. Gochanour, *Chem. Phys. Lett.* **52**, 279 (1977).
- 50 J. Ihm and J. C. Phillips, *Phys. Rev.* **B27**, 7803 (1983).
- 51 T. Tomiki, Y. Ganaha, T. Shikenbaru, T. Futemma, M. Yuri, Y. Aiura, S. Sato, H. Fukutani, H. Kato, T. Miyahara, A. Yonesu and J. Tamashiro, *J. Phys. Soc. Jpn* **62**, 573 (1993).
- 52 G. Lucovsky, R. C. Keezer and E. Burstein, *Solid State Commun.* **5**, 439 (1967).
- 53 H. Tang, F. Lévy, H. Berger and P. E. Schmid, *Phys. Rev.* **B52**, 7771 (1995).
- 54 L. T. Godmanis, A. N. Trukhin and K. Hübner, *phys. stat. sol. (b)* **116**, 279 (1983).
- 55 M. Nagasawa and S. Shionoya, *Solid State Commun.* **7**, 1731 (1969).
- 56 D. Fröhlich, R. Kenkies and Ch. Uihlein, *Proc. Int. Conf. Physics of Semiconductors, Kyoto, 1980, J. Phys. Soc. Jpn* **49**, Suppl. A, 405 (1980).
- 57 T. Miyata and T. Tomiki, *J. Phys. Soc. Jpn* **22**, 209 (1967); *ibid.* **24**, 1286 (1968).
- 58 W. Martienssen, *J. Phys. Chem. Solids* **2**, 257 (1957).

- 59 K. Kobayashi and T. Tomiki, *J. Phys. Soc. Jpn* **16**, 1417 (1961), T. Tomiki, *ibid.* **23**, 1280 (1967).
- 60 A. Matsui, K. Mizuno and M. Iemura, *J. Phys. Soc. Jpn* **51**, 1871 (1982); A. Matsui and K. Mizuno, *ibid.* **51**, 3206 (1982).
- 61 J. D. Wiley, D. Thomas, E. Schönherr and A. Breitschwerdt, *J. Phys. Chem. Solids* **41**, 801 (1980).
- 62 T. Yao and I. Imai, *Solid State Commun.* **9**, 205 (1971).
- 63 J. Nakada, *J. Phys. Soc. Jpn* **20**, 346 (1965).
- 64 D. Dutton, *Phys. Rev.* **112**, 785 (1958).
- 65 Yu. P. Gnatenko, Cand. dissertation, Kiev (1970).
- 66 C. Konak, J. Dillinger and V. Prosser, II–VI *Semiconducting Compounds*, D. G. Thomas, ed. (Benjamin, Inc., New York, 1967), p. 850.
- 67 S. Tutihasi, G. G. Roberts, R. C. Keezer and R. E. Drews, *Phys. Rev.* **177**, 1143 (1969).
- 68 H. Kanzaki, S. Sakuragi and K. Sakamoto, *Solid State Commun.* **9**, 999 (1971).
- 69 H. Zinngrebe, *Z. Phys.* **154**, 495 (1959).
- 70 J. Takeda, T. Ishihara and T. Goto, *Solid State Commun.* **56**, 101 (1985); see also J. Takeda and T. Goto, *Solid State Commun.* **59**, 209 (1986).
- 71 A. Matsui and H. Nishimura, *J. Phys. Soc. Jpn* **49**, 657 (1980).
- 72 H. Nishimura, T. Yamaoka, K. Mizuno, M. Iemura and A. Matsui, *J. Phys. Soc. Jpn* **53**, 3999 (1984).
- 73 K. Mizuno, M. Furukawa and A. Matsui, *J. Phys. Soc. Jpn* **60**, 2768 (1991); M. Sakurai, M. Furukawa, K. Mizuno and A. Matsui, *ibid.* **61**, 445 (1992).
- 74 B. Segall, *Phys. Rev.* **150**, 734 (1966).
- 75 D. T. F. Marple, *Phys. Rev.* **150**, 728 (1966).
- 76 J. G. Liebler, S. Schmitt-Rink and H. Haug, *J. Luminescence* **34**, 1 (1985); J. Liebler and H. Haug, *Europhys. Lett.* **14** (1), 71 (1991).
- 77 H. Tang, H. Berger, P. E. Schmid and F. Levy, *Solid State Commun.* **87**, 847 (1993).
- 78 I. Sildos, A. Suisalu, J. Aarik, T. Sekiya and S. Kurita, *J. Luminescence* **87–89**, 290 (2000).
- 79 M. Watanabe, S. Sasaki, H. Yagasaki and T. Hayashi, *Proc. Int. Conf. on Excitonic Processes in Condensed Matter*, 2001, p. 429.
- 80 L. Grabner, S. E. Stokowski and W. S. Brouer, *Phys. Rev.* **B2**, 590 (1970).
- 81 R. E. Merrifield, *J. Chem. Phys.* **40**, 445 (1964).
- 82 E. I. Rashba, *Soviet Physics-JETP* **27**, 292 (1968).
- 83 A. S. Davydov and A. A. Serikov, *phys. status solidi (b)* **44**, 127 (1971); **42**, 603 (1970).
- 84 K. Cho and Y. Toyozawa, *J. Phys. Soc. Jpn* **30**, 1555 (1971).
- 85 H. B. Shore and L. M. Sander, *Phys. Rev.* **B7**, 4537 (1973).
- 86 H. Sumi, *J. Phys. Soc. Jpn* **36**, 770 (1974); *ibid.* **38**, 825 (1975).
- 87 Y. Toyozawa, *Prog. Theor. Phys. Suppl.* No. 12, 93 (1959).
- 88 U. Heim and P. Wiesner, *Phys. Rev. Lett.* **30**, 1205 (1973); P. Wiesner and U. Heim, *Phys. Rev.* **B11**, 3071 (1975).
- 89 H. Sumi, *J. Phys. Soc. Jpn* **41**, 526 (1976).
- 90 F. Askary and P. Y. Yu, *Phys. Rev.* **B31**, 6643 (1985); see also *ibid.* **28**, 6165 (1983).
- 91 Y. Oka, K. Nakamura and H. Fujisaki, *Phys. Rev. Lett.* **57**, 2857 (1986).

Chapter 11

- 1 As for the mechanism of laser oscillations, the following books are recommended. M. Sargent III, M. O. Scully and W. E. Lamb, Jr, *Laser Physics* (Addison Wesley, 1974); H. Haken, *Laser Theory, Encyclopedia of Physics*, Vol. XXV/2c, S. Flügge, ed. (Springer, 1970).

- 2 L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995) provides the most extensive description of the principles.
- 3 As for nonlinear optics in general, see the following books. Y. R. Shen, *The Principles of Nonlinear Optics* (John Wiley & Sons, 1984); M. D. Levinson and S. S. Kano, *Introduction to Nonlinear Laser Spectroscopy* (Academic Press, 1988).
- 4 D. Fröhlich, *Comments on Solid State Physics* **4**, 179 (1972).
- 5 M. Goepfert-Mayer: *Ann. Physik* **9**, 273 (1931).
- 6 W. Kaiser and C. G. B. Garret, *Phys. Rev. Lett.* **7**, 229 (1961).
- 7 R. Braunstein, *Phys. Rev.* **125**, 475 (1962).
- 8 R. Loudon, *Proc. Phys. Soc.* **80**, 952 (1962).
- 9 J. J. Hopfield, J. M. Worlock and K. Park, *Phys. Rev. Lett.* **11**, 414 (1963).
- 10 M. Kahm and Ch. Uihlein, *phys. stat. solidi (b)* **87**, 575 (1978).
- 11 D. Fröhlich, R. Kenklies and Ch. Uihlein, *J. Phys. Soc. Jpn* **49**, Suppl. A, 405 (1980).
- 12 D. Fröhlich, E. Mohler and P. Wiesner, *Phys. Rev. Lett.* **26**, 554 (1971).
- 13 See for instance, T. Inui, Y. Tanabe and Y. Onodera, *Group Theory and Its Applications in Physics* (Springer, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hongkong, Barcelona, 1990).
- 14 M. Inoue and Y. Toyozawa, *J. Phys. Soc. Jpn* **20**, 363 (1965).
- 15 T. R. Bader and A. Gold, *Phys. Rev.* **171**, 997 (1968).
- 16 M. M. Denisov and V. P. Makarov, *J. Phys.* **C5**, 2651 (1972).
- 17 M. Matsuoka, *J. Phys. Soc. Jpn* **23**, 22 (1967); see also, M. Matsuoka and T. Yajima, *Phys. Lett.* **23**, 20 (1966).
- 18 D. Fröhlich, B. Stagginnus and S. Thurm, *phys. stat. solidi*, **40**, 287 (1970).
- 19 K. Kobayashi, Excitons in thallous halides, in *Excitonic Processes in Solids*, M. Ueta, H. Kanzaki, K. Kobayashi, Y. Toyozawa and E. Hanamura (Springer, Berlin, Heidelberg, New York, Tokyo, 1986), Chapter 7.
- 20 D. H. Fröhlich, *Proc. 10th Int. Conf. on Physics of Semiconductors, 1970, Cambridge, Massachusetts* (US Atomic Energy Commission, 1970) p. 95.
- 21 D. Fröhlich, R. Kenklies, Ch. Uihlein and C. Schwab, *Phys. Rev. Lett.* **43**, 1260 (1979).
- 22 Ch. Uihlein, D. Fröhlich and R. Kenklies, *Phys. Rev.* **B23**, 2731 (1981).
- 23 D. Fröhlich, M. Itoh and Ch. Pahlke-Lerch, *Phys. Rev. Lett.* **72**, 1001 (1994).
- 24 K. Cho. S. Suga and W. Dreybrot, F. Willmann, *Phys. Rev.* **B11**, 1512 (1975).
- 25 M. Itoh and K. Tanimura, *Phys. Rev.* **B53**, 13357 (1996).
- 26 D. Fröhlich, Two and three photon spectroscopy in solids, in *Proc. Todai 1995 and 5th ISSP Int. Symp. on Frontiers in Laser Physics and Spectroscopy*, S. Watanabe, ed. (Pergamon, 1996), p. 29.
- 27 M. Fiebig, D. Fröhlich and K. Kohn, St. Leute, Th. Lottermoser, V. V. Pavlov, R. V. Pisarev, *Phys. Rev. Lett.* **84**, 5620 (2000).
- 28 A. Smekal, *Naturwiss.* **11**, 873 (1923).
- 29 C. V. Raman, *Indian J. Phys.* **2**, 387 (1928).
- 30 R. Loudon, *Adv. Phys.* **13**, 423 (1964); **14**, 621 (1965).
- 31 J. L. Birman, in *Encyclopedia of Physics*, Vol. XXV/2b (Springer, Berlin, Heidelberg, New York, 1974); J. L. Birman and D. N. Pattanyak, in *Light Scattering in Solids*, J. L. Birman, H. Z. Cummins and K. K. Rebane, eds. (Plenum Press, New York, 1979).
- 32 W. Heitler, *The Quantum Theory of Radiation*, 3rd edn (Clarendon Press, Oxford, 1954).
- 33 K. Rebane, V. Hizhnyakov and I. Tehver, *Izv. AN ESSR, Fiz. Mat.* **16**, 207 (1967). This article includes a review of earlier works.
- 34 V. Hizhnyakov and I. Tehver, *phys. stat. solidi* **21**, 755 (1967); *ibid.* **30**, 67 (1970).

- 35 D. L. Huber, *Phys. Rev.* **158**, 843 (1967); **170**, 418 (1968); **178**, 93 (1969).
- 36 E. D. Trifonov and A. S. Troshin, *Physics of Impurity Centers in Crystals* (Tallin, 1972), p. 565.
- 37 P. Y. Yu, Y. R. Shen, Y. Petroff and L. M. Falikov, *Phys. Rev. Lett.* **30**, 283 (1973).
- 38 M. V. Klein, *Phys. Rev.* **B8**, 919 (1973).
- 39 Y. R. Shen, *Phys. Rev.* **B9**, 622 (1974).
- 40 J. R. Solin and H. Merkelo, *Phys. Rev.* **B12**, 624 (1975).
- 41 S. Mukamel, A. Ben-Reuven and J. Jortner, *Phys. Rev.* **A12**, 947 (1975).
- 42 R. Kubo, T. Takagahara and E. Hanamura, in *Physics of Highly Excited States in Solids*, M. Ueta and Y. Nishina, eds., *Lecture Note in Physics*, Vol. 57 (Springer, Berlin, Heidelberg, New York, 1976), p. 304; R. Kubo, T. Takagahara and E. Hanamura, *Solid State Commun.* **32**, 1 (1979).
- 43 Y. Toyozawa, *J. Phys. Soc. Jpn* **41**, 400 (1976).
- 44 A. Kotani and Y. Toyozawa, *J. Phys. Soc. Jpn* **41**, 1699 (1976).
- 45 Y. Toyozawa, A. Kotani and A. Sumi, *J. Phys. Soc. Jpn* **42**, 1495 (1977).
- 46 T. Takagahara, E. Hanamura and R. Kubo; *J. Phys. Soc. Jpn* **43**, 802; 811; 1522 (1977); *ibid*, **44**, 728, 742 (1978).
- 47 M. Aihara, *J. Phys. Soc. Jpn* **48**, 773 (1980); M. Aihara, *Phys. Rev. Lett.* **57**, 463 (1986).
- 48 M. Cardona and G. Güntherodt (eds.), *Light Scattering in Solids, I* (Springer, Berlin, Heidelberg, New York, 1975); *II* (1982); *III* (1982); *IV* (1984); *V* (1989); *VI* (1991).
- 49 M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Clarendon Press, Oxford, 1954); M. Born and M. Bradburn, *Proc. Roy. Soc. A* **188**, 161 (1947).
- 50 R. Kubo, *J. Phys. Soc. Jpn* **17**, 1100 (1962).
- 51 K. K. Rebane, I. Y. Tehver and V. V. Hizhnyakov, *Proc. 1st Soviet-American Symposium on the Theory of Light Scattering in Solids* (Nauka, Moscow, 1978) p. 467.
- 52 A. Kotani, *J. Phys. Soc. Jpn* **44**, 965 (1978).
- 53 P. Nozieres and E. Abrahams, *Phys. Rev.* **B10**, 3099 (1974).
- 54 T. Goto, T. Takahashi and M. Ueta, *J. Phys. Soc. Jpn* **24**, 314 (1968).
- 55 S. Ppermogorov, *phys. stat. solidi (b)* **68**, 9 (1975).
- 56 A. Nakamura and C. Weisbuch, *Solid State Commun.* **32**, 301 (1979).
- 57 H. Kurita, O. Sakai and A. Kotani, *J. Phys. Soc. Jpn* **49**, 1920 (1980).
- 58 Y. Oka and M. Cardona, *Solid State Commun.* **30**, 447 (1979).
- 59 C. Weisbuch and R. G. Ulbrich, *Proc. Int. Conf. on Lattice Dynamics, Paris 1977*, M. Balkanski, ed. (Flammarion, Paris, 1978), p. 167.
- 60 C. Weisbuch and R. G. Ulbrich, *Resonant Light Scattering Mediated by Excitonic Polaritons in Semiconductors*, in Vol. III of Ref. [48], p. 207.
- 61 K. Arimoto, M. Sugisaki, K. Nakamura, K. Tanaka and T. Suemoto, *J. Phys. Soc. Jpn* **63**, 4249 (1994).
- 62 M. Sugisaki, M. Nishikawa, O. Arimoto, K. Nakamura, K. Tanaka and T. Suemoto, *J. Phys. Soc. Jpn* **64**, 3506 (1995).
- 63 M. Sugisaki, O. Arimoto and K. Nakamura, *J. Phys. Soc. Jpn* **65**, 23 (1996).
- 64 K. Nakamura, M. Shigoku, K. Kondo, H. Nagasaki and O. Arimoto, *J. Phys. Soc. Jpn* **70**, 1145 (2001).
- 65 J. Engbring, D. Fröhlich, R. Schepe, S. Spitzer, O. Arimoto and K. Nakamura, *phys. stat. solidi (b)* **196**, 461 (1996).
- 66 W. von der Osten, *Light Scattering in Silver Halides*, in *Light Scattering in Solids, VI* M. Cardona and G. Güntherodt, eds. (Springer, Berlin, Heidelberg, New York, 1991), p. 361, and the references therein.
- 67 W. von der Osten, J. Weber and G. Schaack, *Solid State Commun.* **15**, 1561 (1974).
- 68 J. Weber and W. von der Osten, *Z. Phys. B* **24**, 343 (1976).

- 69 J. Windscheif and W. von der Osten, *J. Phys. C* **13**, 6299 (1980).
- 70 J. Windscheif, H. Stolz and W. von der Osten, *Solid State Commun.* **28**, 911 (1978).
- 71 U. Sliwczuk, H. Stolz and W. von der Osten, *phys. stat. solidi (b)* **122**, 203 (1984).
- 72 K. Nakamura and W. von der Osten, *J. Phys. C* **16**, 6669 (1983).
- 73 W. Wassmuth, H. Stolz and W. von der Osten, *J. Phys. Condensed Matter* **2**, 919 (1990).
- 74 H. Stolz, W. Wassmuth, W. von der Osten and Ch. Uihlein, *Physica* **117B & 118B**, 383 (1983); *J. Phys. C* **16**, 955 (1983).
- 75 H. Stolz, E. Schreiber and W. von der Osten, *In Proc. 17th Int. Conf. Phys. Semiconductors, San Francisco, USA, 1984*, J. D. Chadi and W. A. Harrison, eds. (Springer, New York, 1985), p. 1271; H. Stolz and W. von der Osten, *Cryst. Latt. Def. Amorph. Mat.* **12**, 293 (1985).
- 76 V. Langer, H. Stolz and W. von der Osten, *Phys. Rev. Lett.* **64**, 854 (1990); *J. Luminescence* **45**, 406 (1990).
- 77 L. Van Hove, *Physica* **21**, 901 (1955); **22**, 343 (1956).
- 78 A. Mysyrowicz, J. B. Grun, A. Bivas, R. Levy and S. Nikitine, *Phys. Lett.* **26A**, 615 (1968).
- 79 H. Souma, T. Goto, T. Ohta and M. Ueta, *J. Phys. Soc. Jpn* **29**, 697 (1970).
- 80 O. Akimoto and E. Hanamura, *J. Phys. Soc. Jpn* **33**, 1537 (1972).
- 81 W. F. Brinkmann, T. M. Rice and B. Bells, *Phys. Rev.* **B8**, 1570 (1973).
- 82 J. Usukura, Y. Suzuki and K. Varga, *Phys. Rev.* **B59**, 5652 (1999).
- 83 E. Hanamura, Theoretical aspects of excitonic molecules, in *Excitonic Processes in Solids*, M. Ueta, H. Kanzaki, K. Kobayashi, Y. Toyozawa and E. Hanamura (Springer, Berlin, Heidelberg, New York, Tokyo, 1986), Chapter 2.
- 84 Hanamura, *Solid State Commun.* **12**, 951 (1973).
- 85 E. I. Rashba, Giant oscillator strengths inherent in exciton complexes, in *Exciton at High Density*, H. Haken and S. Nikitine, eds., *Springer Tracts Mod. Phys.*, Vol. 73 (Springer, Berlin, Heidelberg 1975), p. 150.
- 86 G. M. Gale and A. Mysyrowicz, *Phys. Rev. Lett.* **32**, 727 (1974).
- 87 Y. Nozue, T. Itoh and M. Ueta, *J. Phys. Soc. Jpn* **44**, 1305 (1978).
- 88 N. Nagasawa, S. Koizumi, T. Mita and M. Ueta: *J. Luminescence* **12/13**, 587 (1976).
- 89 M. Ueta, T. Mita and T. Itoh, *Solid State Commun.* **32**, 43 (1979).
- 90 M. Inoue and E. Hanamura, *J. Phys. Soc. Jpn* **41**, 1273 (1976).
- 91 M. Shinada and S. Sugano, *J. Phys. Soc. Jpn* **21**, 1936 (1966).
- 92 R. J. Elliot and R. Loudon, *J. Phys. Chem. Solids* **15**, 196 (1960).
- 93 S. Abe, J. Yu and W. P. Su, *Phys. Rev.* **B45**, 8264 (1992); S. Abe, M. Schreiber, W. P. Su and J. Yu, *ibid.* **B45**, 9432 (1992).
- 94 T. Hasegawa, Y. Iwasa, H. Sunamura, T. Koda, Y. Tokura, H. Tachibana, M. Matsumoto and S. Abe, *Phys. Rev. Lett.* **69**, 668 (1992).
- 95 H. Kishida, H. Matsuzaki, H. Okamoto, T. Manabe, M. Yamashita, Y. Taguchi and Y. Tokura, *Nature* **405**, 929 (2000).

Chapter 12

- 1 D. D. Ivanenko and I. Pomeranchuk, *Phys. Rev.* **65**, 343 (1944).
- 2 J. Schwinger, *Phys. Rev.* **70**, 798 (1946); *ibid.* **75**, 1912 (1949).
- 3 J. D. Jackson, *Classical Electrodynamics* (John Wiley and Sons, New York, London, 1962).
- 4 C. Kunz (ed.), *Synchrotron Radiation, Technique and Applications* (Springer, Berlin, Heidelberg, New York, 1979).

- 5 Y. Toyozawa, in *Conf. Digest of 3rd. Int. Conf. on Vacuum Ultraviolet Radiation Physics* (Phys. Soc. Japan, 1971), 30a A3-2.
- 6 A. Kotani and Y. Toyozawa, Theoretical aspects of inner level spectroscopy, Chapter 4 in Ref. [4].
- 7 U. Fano and J. W. Cooper, *Rev. Mod. Phys.* **40**, 441 (1968).
- 8 J. W. Cooper, *Phys. Rev.* **128**, 681 (1962).
- 9 S. T. Manson and J. W. Cooper, *Phys. Rev.* **165**, 126 (1968).
- 10 M. H. Cohen and V. Heine, *Phys. Rev.* **122**, 1821 (1961).
- 11 R. Haensel, G. Keitel, P. Schreiber and C. Kunz, *Phys. Rev.* **188**, 1375 (1969).
- 12 M. Cardona, R. Haensel, D. W. Lynch and B. Sonntag, *Phys. Rev.* **B2**, 1117 (1970).
- 13 F. C. Brown, C. Gähwiller, H. Fujita, A. B. Kunz, W. Scheifley and N. J. Carrel, *Phys. Rev.* **B2**, 2126 (1970).
- 14 A. B. Kunz and N. O. Lipari, *J. Phys. Chem. Solids* **32**, 1141 (1971).
- 15 U. Rössler, *phys. stat. sol (b)* **45**, 483 (1971).
- 16 F. C. Brown, in *Solid State Physics*, Vol. 29, F. Seitz, D. Turnbull and H. Ehrenreich, eds. (Academic Press, New York 1974), pp. 1–73.
- 17 D. E. Sayers, E. W. Lytle and E. A. Stern, in *Advances in X-Ray Analysis*, B. L. Henke, J. B. Newkirk and G. R. Mallet, eds. (Plenum Press, New York 1970), pp. 248–271.
- 18 D. E. Sayers, E. A. Stern and E. W. Lytle, *Phys. Rev. Lett.* **27**, 1204 (1971).
- 19 A. V. Kolobov, H. Oyanagi, K. Tanaka and Ke. Tanaka, *Phys. Rev.* **B55**, 726 (1997).
- 20 E. W. Lytle, D. E. Sayers and E. B. Moore, Jr, *Appl. Phys. Lett.* **24**, 45 (1974).
- 21 For a review, see W. Bambynek, B. Crasemann, R. W. Fink, H.-U. Freund, H. Mark, C. D. Swift, R. E. Price and P. V. Rao, *Rev. Mod. Phys.* **44**, 716 (1972).
- 22 V. O. Kostroun, M. H. Cohen and B. Crasemann, *Phys. Rev.* **A3**, 533 (1971).
- 23 S. Kubota, M. Itoh, J. Ruan (Gen), S. Sakuragi and S. Hashimoto, *Phys. Rev. Lett.* **60**, 2319 (1988).
- 24 M. Itoh, S. Hashimoto and S. Sakuragi, S. Kubota, *Solid State Commun.* **65**, 523 (1988).
- 25 J. L. Jansons, V. J. Krumins, Z. A. Rachko and J. A. Valbis, *phys. stat. solidi (b)* **144**, 835 (1987).
- 26 J. L. Janson, V. J. Krumins, Z. A. Rachko and J. A. Valbis, *Solid State Commun.* **67**, 183 (1988).
- 27 S. Asaka and M. Itoh, M. Kamada, *Phys. Rev.* **B63**, 081104 (R) (2001).
- 28 Y. Toyozawa, *Proc. 10th Taniguchi Symp. Core Level Spectroscopy in Condensed Systems*, J. Kanamori and A. Kotani, eds. (Springer, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1987), p. 231.
- 29 Y. Kayanuma and A. Kotani, *J. Electron Spectr. Relat. Phenom.* **79**, 219 (1996).
- 30 T. Sekikawa, T. Yamazaki, Y. Nabekawa and S. Watanabe, presented at the Annual Meeting of the Phys. Soc. Japan. Hachioji, March 2001.
- 31 P. Morin and I. Nenner, *Phys. Rev. Lett.* **56**, 1913 (1986).
- 32 M. Simon, P. Morin, P. Lablanquie, M. Lavollée, K. Ueda and N. Kosugi, *Chem. Phys. Lett.* **238**, 42 (1995).
- 33 K. Ueda, K. Ohmori, M. Okunishi, H. Chiba, Y. Shimizu, Y. Sato, T. Hayashi, E. Shigemasa and Y. Yagishita, *Phys. Rev.* **A52**, R1815 (1995).
- 34 K. Ueda, H. Chiba, Y. Sato, T. Hayashi, E. Shigemasa and A. Yagishita, *Phys. Rev.* **A46**, R5 (1992).
- 35 M. Simon, C. Miron, N. Leclercq, P. Morin, K. Ueda, Y. Sato, S. Tanaka and Y. Kayanuma, *Phys. Rev. Lett.* **79**, 3857 (1997).
- 36 S. Tanaka, Y. Kayanuma and K. Ueda, *Phys. Rev.* **A57**, 3437 (1998).
- 37 K. Siegbahn, C. Nordling, A. Fahlman, R. Nordberg, K. Hamrin, J. Hedman, G. Johansson, T. Bergmark, S.-E. Karlsson, I. Lindgren and B. Lindberg, in

- ESCA – Atomic, Molecular and Solid State Structure Studied by Means of Electron Spectroscopy (Almqvist and Wicksells, Uppsala, 1967).
- 38 M. Cardona and L. Ley (eds.), *Photoemission in Solids I. General Principles* (Springer, 1978); *II. Case Studies* (Springer, 1979).
 - 39 B. Feuerbach and R. F. Willis, *J. Phys.* **C9**, 169 (1976).
 - 40 C.-O. Almladh and L. Hedin, Beyond the one-electron model, in *Handbook on Synchrotron Radiation, Vol. 1*, E. E. Koch, ed. (North-Holland, Amsterdam, 1983), p. 607.
 - 41 A. Kotani, Inner shell photoelectron process in solids, in *Handbook on Synchrotron Radiation, Vol. 2*, C. V. Marr, ed. (Elsevier, 1987), p. 611.
 - 42 P. H. Citrin, G. K. Wertheim and Y. Baer, *Phys. Rev.* **B16**, 4256.
 - 43 See, for instance, N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, 1976), Chapters 1 and 2.
 - 44 D. Pines, *Elementary Excitations in Solids* (Benjamin, New York, 1963).
 - 45 P. W. Anderson, *Phys. Rev. Lett.* **18**, 1049 (1967). *Phys. Rev.* **164**, 352 (1967).
 - 46 G. D. Mahan, *Phys. Rev.* **163**, 612 (1967).
 - 47 P. Nozières and C. T. De Dominicis, *Phys. Rev.* **178**, 1097 (1969).
 - 48 J. J. Hopfield, *Comments Solid State Phys.* **2**, 40 (1969).
 - 49 J. Friedel, *Comments Solid State Phys.* **2**, 21 (1969).
 - 50 S. Doniach and M. J. Sunjic, *J. Phys.* **C3**, 285 (1970).
 - 51 J. Friedel, *Nuovo Cimento, Suppl.* **7**, 287 (1958).
 - 52 A. Kato, A. Okiji and Y. Osaka, *Prog. Theor. Phys.* **44**, 487 (1970).
 - 53 A. Yoshimoro and A. Okiji, *Phys. Rev.* **B16**, 3838 (1977).
 - 54 Y. Onodera, *J. Phys. Soc. Jpn* **39**, 1482 (1975).
 - 55 G. D. Mahan, *Phys. Rev.* **B11**, 4814 (1975).
 - 56 T. Sagawa, in *Soft X-Ray Band Spectra and the Electronic Structure of Metals and Materials*, D. J. Fabian, ed. (Academic Press, New York 1968), pp. 29–43.
 - 57 C. Kunz, R. Haensel, G. Keitel, P. Schreiber and B. Sonntag, in *Electronic Density of States*, L. H. Bennett, ed. (US Government Printing Office, Washington, DC, 1971), p. 275.
 - 58 H. Neddermeyer, *Phys. Lett.* **44A**, 181 (1973); *Phys. Rev.* **B13**, 2411 (1976).
 - 59 J. D. Dow, in *Proc. 4th Intern. Conf. on Vacuum Ultraviolet Radiation Physics*, E. E. Koch, R. Haensel and C. Kunz, eds. (Pergamon Vieweg, Braunschweig, 1974), pp. 649–661.
 - 60 R. P. Gupta and A. J. Freeman, *Phys. Rev. Lett.* **36**, 1194 (1976).
 - 61 A. Kotani and Y. Toyozawa, *J. Phys. Soc. Jpn* **35**, 1073, 1082 (1973).
 - 62 U. Fano, *Phys. Rev.* **124**, 1866 (1961).
 - 63 S. Suzuki, I. Nagakura, T. Ishii, T. Satoh and T. Sagawa, *Phys. Lett.* **41A**, 95 (1972).
 - 64 A. Kotani and Y. Toyozawa, *J. Phys. Soc. Jpn* **37**, 912 (1974).
 - 65 A. Kotani and Y. Toyozawa, *J. Phys. Soc. Jpn* **37**, 563 (1974).
 - 66 A. Kotani, Inner shell photoelectron processes in solids, in *Handbook on Synchrotron Radiation, Vol. 2*, G. V. Marr, ed. (Elsevier B. V., 1987), Chapter 9.
 - 67 A. Kotani, T. Jo and J. C. Parlebas, Many-body effects in core level spectroscopy of rare earth compounds, *Adv. Phys.* **37**, 37–85 (1988).
 - 68 A. Kotani and S. Shin, *Rev. Mod. Phys.* **73**, 203 (2001).
 - 69 W. Eberhardt, T. K. Sham, R. Carr, S. Krummacher, M. Strongin, S. L. Weng and D. Wesner, *Phys. Rev. Lett.* **50**, 1038 (1983).
 - 70 Kenichiro Tanaka, E. O. Sako, E. Ikenaga, K. Isari, S. A. Sardar, S. Wada, T. Sekitani, K. Mase and N. Ueno, *J. Electron Spectr. and Rel. Phenom.* **119**, 225 (2001).

Chapter 13

- 1 Y. Toyozawa, Fifty years' history of spectroscopic study of condensed matter (in Japanese), *Butsuri*, **51**, 75 (1996).
- 2 K. S. Song and Y. Kayanuma, *Butsuri*, **45**, 469 (1990).
- 3 D. Block, A. Wasiela and Y. M. D'Aubigné, *J. Phys.* **C11**, 4201 (1978).
- 4 C. H. Leung, G. Brunet and K. S. Song, *J. Phys.* **C18**, 4459 (1985).
- 5 Y. Kondo, M. Hirai and M. Ueta, *J. Phys. Soc. Jpn*, **33**, 151 (1972).
- 6 N. Itoh and M. Saidoh, *J. Phys. (France)* **34**, C9–C101 (1973). In this connection, see also review articles and books on defect processes in non-metallic solids: N. Itoh, *Adv. Phys.* **31**, 491 (1982); N. Itoh (ed.), *Defect Processes Induced by Electronic Excitation in Insulators* (World Scientific, Singapore, New Jersey, London, Hongkong, 1989); W. B. Fowler and N. Itoh (eds.), *Atomic Processes Induced by Electronic Excitation in Non-Metallic Solids* (World Scientific, Singapore, New Jersey, London, Hongkong, 1990).
- 7 Y. Toyozawa, *J. Phys. Soc. Jpn* **44**, 482 (1978).
- 8 K. S. Song and C. H. Leung *J. Phys. Soc. Jpn* **56**, 2113 (1987).
- 9 Y. Toyozawa, *J. Phys. Soc. Jpn* **58**, 2626 (1989).
- 10 D. L. Dexter, C. C. Klick and G. A. Russel, *Phys. Rev.* **100**, 603 (1955).
- 11 L. D. Landau, *Phys. Z. Soviet Union*, **1**, 88 (1932); C. Zener, *Proc. Roy. Soc. (London)*, **A 137**, 696 (1932).
- 12 R. Kubo and Y. Toyozawa, *Prog. Theor. Phys.* **13**, 160 (1955).
- 13 Y. Kayanuma, *J. Phys. Soc. Jpn* **57**, 292 (1988).
- 14 H. Sumi, *Phys. Rev.* **B29**, 4616 (1984).
- 15 A. Matui, K. Mizuno, N. Tamai and I. Yamazaki, *Chem. Phys.* **113**, 111 (1987); K. Mizuno, A. Matusi and G. J. Sloan, *Chem. Phys.* **131**, 423 (1989).
- 16 K. Kan'no, M. Itoh and Y. Nakai, *J. Phys. Soc. Jpn* **47**, 915 (1979); O. Arimoto, K. Sakai, H. Nakatani, K. Kan'no and Y. Nakai, *J. Luminescence* **31/32**, 333 (1984).
- 17 A. V. Kolobov and K. Tanaka, Photoinduced phenomena in amorphous chalcogenides: from phenomenology to nanoscale, Chap. 2 in *Handbook of Advanced Electronic and Photonic Materials and Devices, Vol. 5, Chalcogenide Glasses and Sol–Gel Materials* H. S. Nalwa, ed. (Academic Press, San Diego, San Francisco, New York, Boston, Sydney, Tokyo, 2001).
- 18 R. A. Street, T. M. Searle and I. G. Austin, *Phil. Mag.* **29**, 1157 (1974); S. G. Bishop and D. L. Mitchell, *Phys. Rev.* **B8**, 5656 (1973).
- 19 S. J. Berkes, S. W. Ing. and W. J. Hillegas, *J. Appl. Phys.* **42**, 4908 (1971); S. A. Keneman, *Appl. Phys. Lett.* **19**, 205 (1971); J. P. De Neufville, S. C. Moss and S. R. Ovshinsky, *J. Non-Cryst. Solids* **13**, 191 (1973–4).
- 20 K. Kimura, H. Nakata, K. Murayama and T. Ninomiya, *Solid State Commun.* **40**, 551 (1981).
- 21 K. Tanaka, *J. Non-Cryst. Solids* **35/36**, 1023 (1980); K. Tanaka and H. Hisakuni, *J. Non-Cryst. Solids* **198/200**, 714 (1996); K. Tanaka, *J. Appl. Phys.* **24**, 779 (1986).
- 22 K. Tanaka, H. Hamanaka and S. Iijima, in *Proc. 7 Int. Conf. on Amorph. and Liquid Semicond.*, W. W. Spear, ed. (CICL, Univ. Edinburgh, 1977), p. 787; V. F. Krasnov and V. G. Remesnik, *Avtometriia* **4**, 101 (1980); A. V. Kolobov, B. T. Kolomiets, O. V. Konstantinov and V. M. Liubin, *J. Non-Cryst. Solids* **45**, 335 (1981).
- 23 A. V. Kolobov, H. Oyanagi, K. Tanaka and Ke. Tanaka, *Phys. Rev.* **B55**, 726 (1997).
- 24 G. S. Kumar and D. C. Neckers, *Chem. Phys.* **89**, 1915 (1989).
- 25 J. B. Torrance, J. E. Vanzquez, J. J. Mayerle and V. Y. Lee, *Phys. Rev. Lett.* **46**, 253 (1981); J. B. Torrance, A. Giraldo, J. J. Mayerle, J. I. Crowley, V. Y. Lee and P. Batail, *Phys. Rev. Lett.* **47**, 1747 (1981).

- 26 Y. Tokura, T. Koda, T. Mitani and G. Saito, *Solid State Commun.* **43**, 757 (1982).
- 27 Y. Kanai, M. Tani, M. Kagoshima, S. Tokura and T. Koda, *Synth. Met.* **10**, 157 (1984).
- 28 M. Le Conte, M. H. Lemée-Cailleur, H. Cailleur, B. Toudic, L. Toupet, G. Heger, F. Moussa, P. Schweiss, K. H. Kraft and N. Karl, *Phys. Rev.* **B51**, 3374 (1995).
- 29 T. Iizuka-Sakano and Y. Toyozawa, *J. Phys. Soc. Jpn* **65**, 671 (1996).
- 30 A. Moréac, PhD thesis, Univ. of Rennes I, France (1995).
- 31 Y. Nogami, M. Taoda, K. Oshima, S. Koshihara, K. Akimoto, S. Aoki, T. Nakayama and A. Miura, (unpublished).
- 32 A. Painelli and A. Girlando, *J. Chem. Phys.* **87**, 1705 (1987).
- 33 T. Mitani, G. Saito, Y. Tokura and T. Koda, *Phys. Rev. Lett.* **53**, 482 (1984).
- 34 Y. Iwasa, T. Koda, S. Koshihara, Y. Tokura and G. Saito, *Mol. Cryst. Liq. Cryst.* **216**, 195 (1992).
- 35 S. Koshihara, Y. Tokura, T. Mitani, G. Saito and T. Koda, *Phys. Rev.* **B42**, 6853 (1990).
- 36 Y. Tokura, T. Koda, G. Saito and T. Mitani, *J. Phys. Soc. Jpn* **53**, 4445 (1984).
- 37 S. Koshihara, Y. Takahashi, H. Sakai, Y. Tokura and T. Luty, *J. Phys. Chem. B* **103**, 2592 (1999).
- 38 K. Nasu (ed.), *Relaxations of Excited States and Photo-Induced Structural Phase Transitions, Proc. 19 Taniguchi Symposium, Kashikojima, 1996* (Springer, Berlin, Heidelberg, New York, Barcelona, Budapest, Hongkong, London, Milan, Paris, Santa Clara, Singapore, Tokyo, 1997).
- 39 T. Suzuki, T. Sakamaki, K. Tanimura, S. Koshihara and Y. Tokura, *Phys. Rev.* **B60**, 6191 (1999).
- 40 N. Nagaosa, *J. Phys. Soc. Jpn* **55**, 3488 (1986).
- 41 Y. Toyozawa, *Solid State Commun.* **84**, 255 (1992).
- 42 M. Kuwata-Gonokami, N. Peyghambarian, K. Meissner, B. Flugel, Y. Sato, R. Shimano, S. Mazumdar, F. Guo, T. Tokihiro, H. Ezaki, K. Ema and E. Hanamura, *Nature*, **367**, 47 (1994).
- 43 K. Nasu, P. Huai and H. Mizouchi, Photoinduced structural phase transitions and their dynamics, topical review, *J. Phys.: Condens. Matter* **13**, R693–R721 (2001). See also, P. Huai, H. Zheng and K. Nasu, *J. Phys. Soc. Jpn* **69**, 1788 (2000); N. Tomita and K. Nasu, *Phys. Rev.* **B63**, 085 107 (2001); Mizouchi and K. Nasu, *J. Phys. Soc. Jpn* **69**, 1543 (2000); **70**, 2175 (2000).
- 44 Y. Toyozawa, *Acta Physica Polonica A* **87**, 47 (1995).
- 45 J. Hubbard, *Proc. Roy. Soc. (London) A* **276**, 238 (1963); **277**, 237 (1964); **281**, 401 (1964); J. Kanamori, *Prog. Theor. Phys.* **30**, 275 (1963).
- 46 Y. Toyozawa, *J. Phys. Soc. Jpn* **50**, 1861 (1981).
- 47 J. Takimoto and Y. Toyozawa, *J. Phys. Soc.* **Jpn **52**, 4331 (1983).**
- 48 Y. Toyozawa, in *Highlights of Condensed Matter Theory*, F. Bassani, F. Fumi and M. P. Tosi, eds. (North Holland, 1985), pp. 798–830.
- 49 See, for example, L. Salem, *The Molecular Orbital Theory of Conjugated Systems* (Benjamin, New York 1966), p. 110.
- 50 M. Schreiber, *J. Phys. Soc. Jpn* **56**, 1029 (1987).
- 51 R. E. Peierls, *Quantum Theory of Solids* (Clarendon Press, Oxford, 1955), p. 108.
- 52 I. Egri, *Solid State Commun.* **22**, 281 (1971).
- 53 K. Nasu and Y. Toyozawa, *J. Phys. Soc. Jpn* **51**, 2098; 3111 (1982).
- 54 H. Tanino and K. Kobayashi, *J. Phys. Soc. Jpn* **52**, 1446 (1983).
- 55 S. Tomimoto, S. Saito, T. Suemoto, K. Sakata, J. Takeda and S. Kurita, *Phys. Rev.* **B60**, 7961 (1999).
- 56 A. Sugita, T. Saito, H. Kano, M. Yamashita and T. Kobayashi, *Phys. Rev. Lett.* **86**, 2158 (2001).

- 57 Y. Kayanuma, *Phys. Rev.* **B41**, 3360 (1990).
- 58 Y. Kayanuma, *Phys. Rev.* **B62**, 12838 (2000).
- 59 K. Ishida, *Z. Physik, B* **102**, 483 (1997).
- 60 S. Tanaka and Y. Kayanuma, *J. Luminescence* **87–89**, 936 (2000).
- 61 M. Yamashita, T. Manabe, T. Kawashima, and H. Shimizu, *Trends in Chemical Physics* **7**, 231 (1999); M. Yamashita, T. Manabe, T. Kawashima, H. Okamoto and H. Kitagawa, *Coordination Chemistry Reviews* **190–192**, 309 (1999).
- 62 H. Okamoto and M. Yamashita, *Bull. Chem. Soc. Jpn* **71**, 2023 (1998).
- 63 Y. Shinozuka, *J. Phys. Soc. Jpn* **56**, 4477 (1987).
- 64 M. El-Maghraby and Y. Shinozuka, *J. Phys. Soc. Jpn* **67**, 3524 (1998).
- 65 Y. Toyozawa, *Solid-State Electronics* **21**, 1313 (1978); Y. Toyozawa, *Physica* **116B**, 7–17 (1983).
- 66 Y. Shinozuka and Y. Toyozawa, *J. Phys. Soc. Jpn* **46**, 505 (1979).
- 67 K. Takahei and K. Kobayashi, *J. Phys. Soc. Jpn* **43**, 891 (1977).
- 68 D. Lee, A. Mysyrowicz and A. V. Nurmikko, *Phys. Rev. Lett.* **58**, 1987 (1987).
- 69 J. M. Langer, *Inst. Phys. Conf. Ser. No. 135*, Chapter 6. Paper presented at the *Defect Recognition and Image Processing in Semiconductors and Devices Conference, Santander, 1993*.
- 70 D. J. Chadi and K. J. Chang, *Phys. Rev. Lett.* **60**, 2187 (1988); D. J. Chadi and K. J. Chang, *Phys. Rev.* **B39**, 10063 (1989).
- 71 J. Dabrowski and M. Scheffler *Phys. Rev. Lett.* **60**, 2183 (1988).
- 72 M. Kaminska and E. R. Weber, EL2 defect in GaAs, in *Semiconductors and Semimetals, Vol. 38* (Academic Press, Inc., 1993), Chapter 2.
- 73 M. Tajima, *Jpn J. Appl. Phys.* **26**, L885 (1987); M. Tajima, T. Iino and K. Ishida, *Jpn J. Appl. Phys.* **26**, L1060 (1987).
- 74 D. V. Lang, R. A. Logan and M. Jaros, *Phys. Rev.* **B19**, 1015 (1979); D. V. Lang, *J. Phys. Soc. Jpn* **49**, Suppl. A, 215 (1980).
- 75 S. B. Zhang, S.-H. Wei and A. Zunger, *Phys. Rev. Lett.* **84**, 1232 (2000).
- 76 J. M. Langer, *J. Phys. Soc. Jpn* **49**, Suppl. A, 207 (1980).
- 77 J. Nissilä, K. Saarinen, P. Hautojärvi, A. Suchocki and J. M. Langer, *Phys. Rev. Lett.* **82**, 3276 (1999).
- 78 C. H. Park and D. J. Chadi, *Phys. Rev. Lett.* **82**, 113 (1999).
- 79 J. E. Dmochowski, J. M. Langer, Z. Kalinski and W. Jantsch, *Phys. Rev. Lett.* **56**, 1735 (1986).
- 80 N. Itoh and A. M. Stoneham, *Materials Modification by Electronic Excitation* (Cambridge University Press, Cambridge, 2001).

Chapter 14

- 1 D. Voet and J. G. Voet, *Biochemistry* (J. Wiley & Sons, New York, 1995).
- 2 E. A. Newsholme and A. R. Leech, *Biochemistry for Medical Sciences* (J. Wiley, New York, 1983), p. 16.
- 3 Y. Toyozawa, *J. Phys. Soc. Jpn* **69**, 1907 (2000).
- 4 S. L. Miller and L. E. Orgel, *The Origin of Life* (Prentice Hall, Englewood Cliffs, NJ, 1973).
- 5 D. Mauzerall, Oceanic sunlight and origin of life, in *Encyclopedia of Earth System Science, Vol. 3* (Academic Press, Inc., 1992) pp. 445–453.
- 6 P. R. Gast, in *Handbook of Geophysics and Space Environments*, S. L. Valley, ed. (Airforce Cambr. Res. Lab., Office of Aerospace Res. and US Airforce, 1965), p. 16.

- 7 H. Noda, *Origin of Life* (in Japanese) (Baifukan, Tokyo, 1996).
- 8 Ref. [1] p. 22.
- 9 S. L. Miller, *Science* **117**, 528 (1953).
- 10 A. Kawamori, in *Electron and Life* (in Japanese, *New Biophysics Series II, Vol. 1*, Kyoritsu, Tokyo, 2000), Section 2.2; *Biochim. Biophys. Acta* No. 1503, 112 (2001).
- 11 H. Reiche and A. J. Bard, *J. Am. Chem. Soc.* **101**, 3127 (1979); W. W. Dunn *et. al.*, *J. Am. Chem. Soc.* **103**, 6893 (1981).
- 12 E. Imai, H. Honda, K. Hatori, A. Brack and K. Matsuno, *Science* **283**, 831 (1999).
- 13 J. Onoe and T. Kawai, *Chem. Lett.* 1167 (1985).
- 14 R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics* (The University Press, Cambridge, 1939).
- 15 J. Castresana *et. al.*, *EMBO (European Molecular Biology Organization) J.* **13**, 2516 (1994); J. F. Kasting, *Science* **259**, 920 (1993).
- 16 K. Matsuura and S. Itoh, *Kagaku (Iwanami)* **68**, 839 (1998).
- 17 M. Akiyama, in *Shinka (Evolution, in Japanese)* (The University Press, Tokyo, 1991), Vol. 5, p. 49.
- 18 L. Margulis and R. Fester (eds.), *Symbiosis as a Source of Evolutionary Innovation* (MIT Press, Cambridge, 1991).
- 19 L. J. Bjorn, *Ljus och Liv* (Aldus/Bonnier, Stockholm, 1973); German translation: *Photobiologie, Licht und Organismen* (Gustav Fischer, Stuttgart, 1975), Chapter 8.

Appendixes

- 1 G. Wannier, *Phys. Rev.* **52**, 191 (1937).
- 2 Y. Kayanuma, *J. Phys. Soc. Jpn* **57**, 292 (1988).

Subject index

- $A(B)$ coefficient of Einstein, **28–29**
absorption coefficient, **136**
absorption process, 28, **33–37**
absorption spectrum, 30
acceptor, **110**, 335
acceptor state, 110
acoustic mode, 65–67, 72, **77**, 168–169
acoustic polaron, 168–176
active mode (for J-T splitting), 56–57
additional boundary condition, 238
adiabatic limit, **176**, 347
adiabatic potential, **40–42**, 53–64, 183,
alkali halides, 117, 128, 137, 149, 168, 180, 226
allowed band edge type (first class of exciton)
 absorption, **125**, 242, 245
alloy model, 356
alternating mode, 353–354
amalgamation (A) type, 132, 137, 143–148
amino acid, 380, 382, 387
amorphous chalcogen(ide), 334
amplification (of electromagnetic wave), 29
angle-resolved photoelectron spectroscopy, 229
angular frequency, 3
anharmonicity, 87
anion vacancy, 326
anisotropic exchange energy, 246
annihilation, 293
annihilation operator, **20**, 32, 35, 336, **391**
anomalous dispersion, 240
antibonding orbital, 188, 322, 324–325, 368
anticommutation rule, **391**
antiferromagnetic order, 340
anti-neutrino, 106
antisite atom, 359, 368
aromatic molecule, 352
aromatic molecular crystal, 186, 226, 229
asymmetric Lorentzian lineshape, 198–199
atomic orbital, 286, 325
attenuation coefficient, **136**
Auger photoelectron, 298
Auger process, 293–295, 298
Auger width, 295
Auger-free luminescence, 296–297
autotrophs, 389
average t -matrix approximation (ATA), 215
averaged oscillator strength per state (AOSPS), 221,
 225
azimuthal quantum number, 322
bacteria, 388
bandgap, 40, 79, **109–110**, 117, 127, 329, 333
basic lattice vector, **72**
basic reciprocal lattice vector, **72**, 77
biexciton (*see* excitonic molecule)
binding energy of core electron, 299
binding energy of excitonic molecule, 279–281
biological evolution, 382, 389
biomolecule, 372, 382
biopolymer, 380, 383, 387
bistability, 321–322, 335, 359–370
Bloch function (orbital), 107, 110, 118, 350, 392
Bloch theorem, 107, 124
Boltzmann distribution, 7, 24
bond breaking, 298, 368
bonding orbital, 188, 286, 324–325, 368
Born approximation, 102
Born–Oppenheimer approximation, 40–42
boson (Bose particle), 20
bound exciton, 141, 215
“bra” vector, 17, 26
Bremsstrahlung, 284
Brillouin scattering, 274
Brillouin zone (boundary), 73, 118, 277
broken parity, 328
cancellation of oscillator strength, 287
cancellation of potential, 287
canonical equations of motion, 4–6, 31–33
canonical transformation, 156
carbohydrate, 372
Carnot cycle, 389
carrier, 109
catalyst, 371, 387
causality, 89, 93, 389
central cell correction, 112
centrifugal potential, 288–289, 322

- charge density, 2, 100
- charge density wave (CDW), 347, 353, 356
- charge transfer, 335, 347
- charge transfer compound, 335
- charge transfer exciton, 342, 355
- charge transfer insulator, 356
- chemical energy, 372–378, 387–388, 390
- chemical evolution, 375, 378–383, 387–389
- chemical reaction, 285, 373, 385
- chemical shift, 284, 299
- chlorophyll, 373, 378
- chloroplast, 389
- circularly polarized light, 103, 244
- cis* form, 335
- closed shell, 325–326, 335, 352
- c-number, 11, 90
- coal, 388
- coherence, 239
- coherent light, 29
- coherent potential approximation (CPA), **137–138**, 202, 215, 356
- cohesive energy, 87
- color center, 149
- commutation relation, 11, 16, 22, 37, 156, 392
- commutativity, 156
- commutator, **11**, 392
- complementarity, 22
- completeness, 74
- compound semiconductors (III–V), 186, 226, 324, 366–369
- compound semiconductors (II–VI) 226, 229, 273, 274, 281
- Condon approximation, 44, 48, 56
- conduction band, 108–109
- conductivity spectra, 137
- configuration coordinate (C.C.) model, **45–47**, 212, 297, 320, 329, 344, 375
- configuration interaction, 119
- conjugate operator, 12
- conjugate polymer, 283
- conjugate variable, 6
- continuity equation, 2
- continuum model (for self-trapping), 176, 360
- cooperative excitation, 343
- coordination number, 334
- core electron, 127, 284
- core exciton, 79
- core hole, 293, 296
- core hole–valence exciton complex (CHVXC), 303
- corpuscular picture, 133
- correlation effect, 167, 283, 308
- correlation function, 256–257
- correlation of absorption and emission spectra, 213–220, 262–270
- correlation time, 257
- correspondence principle, 19,
- coupled wave, 133, 134
- coupling coefficient, 45–46, 54, 58, 62, 173
- coupling function, 46, 193, 205–206
- coupling strength, 49, 50, 65, 213
- covalent bond, 73, 279, 321–324, 368
- covalent crystal, 108, 324
- creation operator, **20**, 32, 35, 311, **391**
- cross luminescence, 296
- crystal field, 132
- cumulant, 229, 256, 267, 300, 305–306
- cuprous halide, 117, 131, 132, 136–137, 279–282
- cuprous oxide, 117, 272, 280–282
- cyanobacteria, 388–389
- cyclic boundary condition, 5, 12, 70, 73, 93, 108
- cyclotron resonance, 152, 155
- damping, 136
- Darwinism, 388
- Davydov splitting, 116
- de Broglie wavelength, 41
- Debye–Waller factor, 292
- defect, 360–361
- defect formation, 285
- deformation potential, 169, 177, 360
- degeneracy, 29, 54, 128
- degree of asymmetry, 199
- degree of ionicity, 335
- degree of localization, 178
- delocalization, 138
- delocalized exciton, 193
- density matrix, 23–27, 250, 262
- density of states, (DOS), 139, 221
- dephasing, 47
- depolarizing field, 80–83, 96–98, 101, 116, 135
- desorption, 286, 298
- detailed balance, 379
- Dexter–Klick–Russel criterion, 331–332
- diamond structure, 108, 324
- dielectric constant, **2**, 89, 96, 133, 136, 291, 384–385
- dielectric continuum, 117, 168, 177,
- dielectric dispersion, 77, 133–134
- dielectric function, 80
- dielectric tensor, 103, 106
- difference frequency generation, 240
- dilation field, 169, 177, 360
- dilation, 169
- dimerization, 321, 336–338, 341–342
- dipole approximation, 34
- dipole–dipole interaction, 114–115, 122, 333, 342–344
- direct exciton transition, 196, 272
- direction of polarization, 5, 242–243
- disorder-assisted transition, 148
- dispersion 70, 133–134
- dispersion relation (K.–K. relation), **95**, 127, 136–137, 302
- dissociation, 298
- D* method, 96, 98, 123, 149
- donor, **109**, 335
- donor state, 110
- d-subshell, 308
- Dulong–Petit’s law, 384
- DX center, 359, 367
- dynamical coherent potential approximation (d-CPA), 176, 223

- effective charge, 78, 82–83
- effective mass, 150
- effective mass approximation, 110, 122
- effective mass tensor, 111, 276
- eigenfunction, 13
- eigenvalue, 13, 16
- eigenvalue problem, 15, 16
- Einstein–de Broglie relation, 21
- EL2 center, 367, 369
- elastic (Rayleigh) scattering of light, 241, 251–252
- elastic compliance, 71
- elastic continuum, 72, 168, 177
- electric current density, 2
- electric dipole moment, 34
- electric field, 1
- electric flux density, 1
- electric polarization, 133, 153
- electric quadrupole moment, 105
- electric susceptibility, 89, 96, 123
- electrolysis, 385
- electromagnetic field, 1, 133
- electromagnetic wave, 1, 3
- electron affinity, 112, 146, 321
- electron bubble, 187
- electron energy-loss spectra (EELS), 102
- electron gas, 83, 100
- electron nuclear double resonance (ENDOR), 327
- electron spin resonance (ESR), 168, 188
- electron–hole drop, 280
- electronic dielectric constant, 65, 80, 302, 360
- electronic polarization, 80, 162, 302
- electronic polaron, 162–168, 291, 302
- electron–phonon coupling constant, 173
- electron–phonon interaction, 40–68, 149–162, 168–191, 254–279, 326–370
- electrostatic instability, 336–337
- elementary excitation, 10, 69, 113, 118, 132, 192, 284
- elementary process, 135, 247
- E* method, 96–97, 123, 149
- emission process, 28, 33–35
- emission spectrum, 30
- energy band, 107
- energy (potential) barrier for self-trapping, 179, 183, 195, 219, 333, 359–360, 366
- energy density of electromagnetic field, 2, 3
- energy gap, 296, 340, 353
- ergodic theory, 52, 252
- ergodicity, 263, 266
- eukaryotic cell, 389
- evolution, 381
- exchange energy (isotropic), 122
- exchange interaction, 62, 119, 121, 128–132, 318, 394
- excimer, 186, 189, 322, 335
- excimer laser, 323
- excitation spectrum, 265, 272
- exciton, 113–148
- exciton band, 115, 217, 291
- exciton bubble, 188–189
- exciton gap, 329
- exciton polariton, 137
- exciton sideband, 167
- excitonic molecule (biexciton), 279–282
- excitonic string, 342–344
- exciton–LO-phonon bound state (EPBS), 208, 303
- exciton–phonon coupling constant, 219
- exciton–phonon instability, 332
- exciton–phonon interaction, 194–195, 271–278, 296–298
- expectation value, 11, 25, 27
- extended X-ray absorption fine structure (EXAFS), 292, 334
- extrinsic self-trapping, 186, 359–366
- extrinsic semiconductor, 109
- F band, 53
- F center, 53, 326–328, 366
- F' center, 366
- face-centered-cubic (f.c.c.) lattice, 73
- false ground state, 321, 355
- Fano effect, (antiresonance), 193, 199, 310, 316
- fat, 372
- Fermi distribution, 101, 357
- Fermi edge singularity, 307, 310
- Fermi energy, 101, 109, 309, 317, 368
- Fermi wave number, 101, 353
- Fermi's golden rule, 33, 151
- fermion (Fermi particle), 391
- first Brillouin zone, 73, 107, 292
- fluctuation–dissipation theorem, 149, 301
- forbidden band (*see also* bandgap), 110
- forbidden band edge type (second class of exciton) absorption 125, 242, 245
- force range, 168, 177–178
- form factor of exciton–phonon interaction, 194
- Franck–Condon principle, 43, 218, 343, 363
- Franck–Condon state, 45, 210, 319, 323, 375
- Franz–Keldysh effect, 216
- free-electron model of metals (Drude–Sommerfeld theory), 97, 304
- free energy, 25, 48, 343, 375–376, 379, 387
- Frenkel defect, 330, 332
- Frenkel exciton, 113, 138, 333
- frequency, 3
- Friedel sum rule, 306
- frustrated spin, 352
- F–S discontinuity line, 180, 184–185, 362, 365
- f-subshell, 308,
- F-type emission, 177, 183, 219, 226–227, 246
- full width at half-maximum, 136, 200
- Gaussian lineshape, 57, 220
- gene, 389
- generating function, 50, 300, 315
- germanium, 108
- giant band, 287
- giant oscillator strength, 141, 145, 216, 281
- giant two-photon absorption (GTA), 279–281
- glass, 334
- glide symmetry, 76

- hard core, 69
- harmonic oscillation, 6, 19, 71
- harmonic oscillator, 6, 10, 15, 23–24, 75
- Hartree–Fock approximation, 118, 167, 299
- Hartree–Fock equations, 394
- H-center, 326, 328
- Heisenberg equation of motion, 13, 15, 19, 37
- Heisenberg picture, 13, 19, 25, 90
- Heisenberg representation, 90, 252, 256
- Heitler–London theory, 325, 346, 349
- helical structure, 76, 104
- helicity, 105–106
- Hermite operator, 12, 32
- Hermiticity, 13, 90
- heterotrophs, 389
- hexagonal-close-packed (h.c.p.) lattice, 73
- highest occupied molecular orbital (HOMO), 385
- hole-mediated bond, 327
- homochirality, 103
- homogeneity, 90
- hot luminescence, 31, 47, 259, 395
- Hubbard band, 283
- Hubbard Hamiltonian, 346
- Hubbard–Peierls system, 353
- Hückel’s $4n+2$ rule, 352
- Hund–Mulliken theory, 346, 349
- hydration, 385–387
- hydrogen bond, 383–384
- hydrolysis, 385
- hydrothermal vent, 374, 387
- hyper-Raman scattering, 279–282
- impurity band, 145
- impurity center, 53
- impurity cluster, 148
- impurity percolation, 363
- inactive mode (for J-T splitting), 56–57
- incoming resonance, 272
- incomplete (sub)shell, 308–309
- indirect absorption edge, 200
- indirect exciton (transition), 196, 235, 274, 280
- induced emission, **28–29**
- inelastic light scattering, 252
- inelastic X-ray scattering, 318
- infrared active, 81
- infrared divergence, 306, 310
- inner shell, 284–286, 293
- insulator, 109
- interaction mode, 45, 53, 218, 332, 353, 355, 375, 377, 395
- interaction representation, 90
- interatomic bond, 319, 322, 375
- interatomic forces, 69, 319
- interatomic potential, 69
- interband transition, 124
- interbond angle, 326
- interchain coupling, 341
- interference effect between exciton lines, 196
- intermediate coupling theory of polaron, 155
- intermediate state resonance, 241, 273–274
- intervalley scattering, 275
- intramolecular vibration, 209
- intrinsic semiconductor, 109
- inversion symmetry, 76, 240
- ionic (I-)phase, 335
- ionic crystal, 321, 370
- ionization continuum, 117, 162
- ionization energy, 321
- irreducible representation, 54, 62, 243–246
- isoelectronic impurity, 112, 359, 362
- isomer, 103
- Jahn–Teller effect, 54, 352, 369
- Jahn–Teller split spectrum 54–64
- Jahn–Teller theorem, 53, 351
- j – j coupling, 129
- joint density of states, 124
- Joule heat, 2
- “ket” vector, 17, 26
- Koopman’s theorem, 299
- Kramers doublet, 132
- Kramers–Kronig relation (*see* dispersion relation)
- K-shell, 285
- Landé factor, 105
- large polaron, 180
- laser, 29, 239, 296
- lattice constant, 66, 69, 321
- lattice defect, 53, 109, 149, 333, 358
- lattice relaxation, 45, 173, 254, 296, 326, 329, 344, 359
- lattice relaxation energy, 55, 178, 218, 330, 346
- lattice vector, 72, 242
- lattice vibrations, 69–88
- law of refraction, 3
- lead halide, 229
- librational mode, 212
- lifetime broadening (of Lorentzian spectrum), 192, 221, 293
- light scattering, 33, 241, 247–248
- linear operator, 11
- linear response theory, 77, 89
- linearly polarized light, 244
- linewidth, 200, 202, 221, 242, 294
- linked cluster expansion, 313
- lipid, 380
- living activities, 371–372, 378, 383, 387
- localization, 138
- localized electron, 40, 254, 262, 395
- localized excitation, 193, 217
- localized mode, 46, 83, 86
- lone pair, 323, 334
- long-range electron–defect coupling constant, 361, 365
- long-range electron–phonon coupling constant, 179, 361, 365
- long-range electron–phonon interaction, 66, 168, 178, 180
- longitudinal acoustic (LA) phonon, 276
- longitudinal effective mass, 112

- longitudinal exciton, 122, 127, 135, 162, 281
- longitudinal optical mode, 150
- longitudinal optical phonon, 65, 272
- longitudinal susceptibility, 98
- Lorentz force, 2, 31–32, 284
- Lorentzian shape, 220–221, 251
- low-energy electron diffraction (LEED), 293
- lower-branch polariton, 135
- lowest unoccupied molecular orbital (LUMO), 385–386
- L–S coupling, 129
- L-shell, 285
- ℓ -t splitting, 116, 126–128, 136–137
- luminescence, 30, 259, 272
- macroscopic size, 6–4
- Madelung constant, 337
- Madelung potential, 284, 288, 296, 326, 335
- magnetic dipole moment, 105, 246, 322
- magnetic field, 1
- magnetic flux density, 1
- magnetic gap, 341
- magnetic permeability, 2
- matrix element, 18
- matrix mechanics, 10
- Maxwell equations, 1, 17
- mean field theory, 99, 341, 356
- metal, 99
- metallic bond, 324
- metallic conduction, 109
- metastability, 339
- metastable state, 174, 177, 186, 361, 383
- microscopic electric field, 96
- mitochondrion, 389
- mixed crystal, 131, 137, 143, 186, 228, 356
- mixed state, 26
- mobility, 152, 169, 187
- molecular crystal, 111, 209, 233
- molecular dynamics method, 356
- molecular field theory, 99
- momentarily localized exciton, 193, 215, 225
- momentum density of electromagnetic field, 3
- monochromaticity, 239
- monovalent metal, 345
- Monte-Carlo distribution, 220
- most divergent term approximation, 313
- motional narrowing, 193, 200–202, 220, 257
- motional reduction of phonon sideband, 204
- Mott insulator, 283, 356
- M-shell, 285
- multielectron transition, 343
- multiphonon scattering of exciton, 201
- multiphoton process, 239
- multipole expansion, 114
- multistability, 383
- mutation, 289
- natural selection, 382–383, 388–390
- nearly free electron approximation, 107, 292
- negative entropy, 376
- negative U effect, 347, 357, 364
- neutral (N-)phase, 335
- neutral–ionic domain wall (NIDW), 342, 345
- neutral–ionic (N–I) phase transition, 335, 340, 343
- neutrino, 106
- neutron diffraction, 212, 338
- neutron scattering, 338
- non-adiabaticity parameter, 173
- non-crystalline systems, 293
- nonlinear optical response, 239
- non-negative operator, 23
- nonradiative annihilation, 329
- nonradiative de-excitation, 332, 375
- nonradiative transition, 331
- normal dispersion, 240
- normal mode, 44, 70, 72, 75, 77
- normalization condition, 12, 16
- N-shell, 285
- n-type semiconductor, 109
- nuclear motion, 40
- number operator, 17, 336
- old quantum theory, 22
- one-electron picture, 107
- on-site Coulomb repulsion, 345, 356
- on-site electron–phonon interaction, 346
- optical activity, 77, 103
- optical conversion (exciton to biexciton), 281
- optical ionization energy, 361
- optical mode, 65, 67, 72, 77, 133
- optical pumping, 323
- optically active crystal, 103
- orbital angular momentum, 105
- orbital degeneracy, 352
- order parameter, 321, 336
- orthoexciton, 246
- orthogonal (quantum mechanics), 13
- orthogonality catastrophe, 306
- orthogonalization, 287–289
- orthonormal complete set, 5, 23, 108, 110
- orthonormality, 74, 108, 347
- oscillator strength, 37, 130, 132, 292, 379–380
- outer subshell, 293, 308
- outermost core state, 296
- outgoing resonance, 272
- overlap integral, 302
- oxonium, 385
- paraexciton, 246
- partition function, 24
- path integral method, 159
- Pauli exclusion principle, 118, 309, 337
- Peierls instability, 351
- Peierls theorem, 353
- Peierls transition, 347
- periodic structure, 69
- periodic table, 308, 322–324
- persistence (P) type, 132, 137, 143–148
- persistent photocurrent, 359, 364
- phase (operator), 22
- phase-matching condition, 240
- phase shift, 306

- p-hole picture (for covalent bond), 322–326, 368
- phonon, 50, 242
- phonon absorption, 151
- phonon dressed electron, 152
- phonon emission, 151
- phonon field, 152, 360
- phonon polariton, 137
- phonon polaron, 162
- phonon sideband (phonon structure), 51, 193, 204, 207, 209, 229
- phonon vacuum state, 151
- phonon-assisted transition, 148
- phonon-mediated electron–hole interaction, 185, 364–365
- photocatalyst, 386–387
- photochemical reaction (process), 321, 335, 371, 375–379, 382–384, 388–390
- photodarkening, 333
- photoelectron spectroscopy, 298–299, 317
- photoelectron spectrum, 293
- photoinduced phase transition (PIPT), 341
- photoinduced structural change, 319
- photoisomerization, 334
- photopolymerization, 387
- photostructural change, 334, 389
- photosynthesis, 373, 387–389
- pitch of helical structure, 76
- Planck constant, 9
- plasma frequency, 101
- plasma oscillation, 83, 100–101, 304
- plasmon satellite, 304
- point defect, 360
- point group, 243
- point–dipole interaction, 169
- Poisson distribution, 49, 209, 234, 302
- polariton, 128, **132–133**, 194, 236, 242, 274, 282
- polariton bottleneck, 31, 194, 234, 236
- polarizability, 134
- polarizability operator, 252
- polarization, 133, 149
- polaron, 149, 153, 179, 297, 356
- polaron coupling constant, 152
- polaron effect, 118, 193, 364
- polaron mass, 155
- polaron radius, 154, 159, 297
- polymerization of biomolecules, 387
- population inversion, 296, 323, 374
- positive hole, 109
- positive-definite operator, 23
- potentially active mode (for J-T splitting), 56–57
- Poynting vector, **2**, 3
- principal quantum number, 111, 285, 322
- probability amplitude, 37
- proliferation, 342
- propagator, 251
- protein, 372, 383,
- pseudo Jahn-Teller effect, 62, 64, 328–329, 350–352, 368
- pseudo momentum, 150
- pseudo-potential, 286, 290
- pseudo-potential theory, 289
- p-subshell, 285
- p-type semiconductor, 110
- pumping, 29
- pure state, 26
- q-diagonal representation, 11, 21
- q-number, 11
- quantization, 10
- quantum beat, 279
- quantum mechanics, 10–27
- quantum Monte-Carlo (QMC) method, 159
- quantum number, 38
- quantum yield, 331, 333
- quartz, 103, 226
- quasi-localized mode, 83, 86
- Rabin–Klick parameter, 190
- radiation–matter coupling constant, 239
- radiation–matter interaction, 31
- radiative annihilation, 293, 295, 329, 332
- radiative decay (damping), 248, 271, 293
- radiative de-excitation, 376
- radiative width, 250, 295
- Raman active, 81
- Raman scattering, 81, 241, 247, 253
- Raman–antiStokes line, 81, 247, 260, 274
- Raman–Stokes line, 81, 247, 260, 272, 274
- random phase approximation, 90
- rapid modulation, 257
- rare-gas solid, 187–188
- Rayleigh scattering, 247, 251, 253, 257
- reaction coordinate, 375, 377
- real process, 247, 251, 272
- reciprocal lattice vector, 72
- recombination luminescence, 296
- redistribution of oscillator strength, 286
- redox potential, 321, 326, 335
- reduced mass, 117, 122
- reduction, 326
- reflectivity, 136, 236
- reflectivity spectrum, 137
- refractive index, **3**, 103, 133, 136, 239
- relative motion, 118, 147
- relaxation, 173, 216, 320, 375
- relaxation energy, 47, 55, 173, 232, 302, 375–377
- relaxed excited state (RES), 47, 319, 333, 335, 350, 375–376, 379, 381–382
- renormalization effect, 299
- representation (quantum mechanics), 18
- residual dielectric constant, 128, 133
- resolvent operator, 139, 271, 275
- resonance fluorescence, 248
- resonance near threshold, 290
- resonance transfer, 322
- resonant Brillouin scattering, 274
- resonant light scattering, 254, 271
- resonant Raman scattering, 247, 353
- resonant secondary radiation, 31, 247, 253
- resonant two-phonon scattering, 152

- respiration, 388
- retarded potential, 4
- rotational motion, 42
- Rydberg constant, 112
- Rydberg states, 64, 117, 291, 297
- satellite, 285, 301
- scalar potential, 4, 31
- scalar product (quantum mechanics), 12, 17
- scaling argument, 223, 227
- scattering amplitude, 292
- scattering phase shift, 292
- Schrödinger equation, 10, 11, 90, 110, 248, 392
- Schrödinger picture, 13, 19, 25, 90
- screened Coulomb potential, 96, 117, 121, 361
- screening distance, 102
- screening effect, 100, 123, 127, 129, 167, 291, 365, 368, 385
- screw dislocation, 382
- s-d mixing interaction, 311
- second harmonics generation, 240
- second-order susceptibility, 239
- second quantization, 391
- selectivity, 378, 381
- selenium, 226
- self-conjugate operator, 12
- self-consistent field, 158
- self-decomposed exciton, 183, 189, 355
- self-energy, 138, 153, 157, 275, 312, 314
- self-energy matrix, 197
- self-induced potential, 154, 164
- self-interaction, 178
- self-organization, 29
- self-replication, 381, 383
- self-trapped electron (STEL), 188, 190, 324
- self-trapped exciton (STE), 189–190, 207, 320, 326–327, 331–333, 354
- self-trapped hole (STH), 168, 188–190, 323, 326–327, 364
- self-trapped state, 149, 177, 179, 187, 213, 219, 352
- self-trapping, 149, 168, 176, 278, 320, 331, 355–356
- semiconductor, 109
- shallow D state, 359, 363
- shallow–deep bistability, 359–361, 363, 366
- shell (atomic), 285
- shift and width of exciton line, 196
- short-range electron–defect coupling constant, 361–362, 365
- short-range electron–phonon coupling constant, 179, 361–362, 365
- short-range electron–phonon interaction, 66, 168–169, 178, 180
- sideband, 285, 301
- silicon, 108
- silver halides, 186, 226–228, 276,
- simple crystal lattice, 72
- simultaneous excitation, 285, 299, 316
- singlet exciton, 120, 123, 129–130, 274
- singlet–triplet conversion, 274
- site-selective excitation, 285, 318
- Slater determinant, 118, 123, 325, 391
- slow modulation, 257
- small polaron, 180
- snow-ball, 188
- solar radiation, 371–373, 378–379, 384–385, 387, 389
- soliton, 356
- sound wave, 72
- sp hybridized orbital, 108, 324
- spatial dispersion, 106
- specific structure of biomolecules, 382
- spectral density of normal modes, 5, 36
- spectral distribution of radiation, 7
- spin angular momentum, 106
- spin density wave (SDW), 347, 353, 356
- spin soliton, 341
- spin wave, 340
- spin–orbit interaction, 62, 128, 132, 313
- spin–orbit splitting, 129, 131, 245
- spin–orbit-allowed transition, 53
- spontaneous emission, 28–30, **39**
- square well potential, 182
- s-subshell, 285
- static dielectric constant, 65, 80, 83, 118, 360
- stationarity, 90
- stationary spectrum, 395
- stationary state, 19
- steepness coefficient, 214, 223, 226
- steepness index, 223–224
- stochastic model, 254, 259
- stochastic variational method, 280
- Stokes shift, 52, 220, 247, 359, 376
- strain, 71, 168
- stress, 71
- strong correlation, 109
- strong coupling theory of polaron, 158–161
- strong coupling, 217, 219–220, 257, 268
- strong scattering regime, 141, 201
- strong scattering, 219–220
- structural phase transition, 321, 336
- S-type emission, 183, 219, 226–227, 354, 363
- subband, 340
- subgiant band, 287
- subshell (atomic), 285, 308, 322, 326
- substitutional impurity, 359–360
- successive Raman scattering, 259
- sum frequency generation, 240
- sum rule, 37, 79, 95
- superconductivity, 347
- superposition principle, 89
- susceptibility tensor, 103
- symbiosis theory, 389
- synchrotron, 284
- synchrotron radiation (SR), 284
- thallous halides, 186, 226
- thermal equilibrium, 7, 28, 52, 372
- thermal ionization energy, 361
- thermal radiation field, 7, 28
- thermalization, 30–31, 236,
- thermally induced emission, 29
- thermodynamical average, 24, 27

- third-order susceptibility, 241, 283
- tight binding approximation, 107, 216
- time-ordered exponential, 213
- time-resolved spectroscopy, 269, 279, 297, 355–356, 395
- time-reversal symmetry, 132
- titanium dioxide, 226, 229–230
- Tl⁺-like center, 53
- t*-matrix, 140
- total angular momentum, 129
- total reflection, 136–137
- trace, 23, 257
- trans* form, 335
- transfer energy, 138, 209, 216, 232, 236, 345
- transient spectrum, 395
- transition dipole moment, 34, 43, 114, 122, 126
- transition element, 308
- translational mass, 122
- translational motion, 118, 216
- translational symmetry, 74, 114, 170–171
- transport coefficients, 89
- transverse effective mass, 112
- transverse exciton, 122–123, 127, 129, 281
- transverse optical phonon, 150
- transverse susceptibility, 97
- transverse wave, 3, 32
- triplet exciton, 120, 129–130, 274
- TTF-CA, 338, 340–341
- turning point, 211
- T–U–S* model, 345–359
- two modes model for impurity Urbach rule, 215, 227
- two-photon absorption, 241–242, 280–281

- Umklapp process, 151, 171
- uncertainty, 20
- uncertainty principle, 20–21, 154, 192, 395
- uncertainty relation, 6, 20, 22
- uniaxial crystal, 82, 132
- unit cell, 14
- unitary transformation, 35
- unrestricted Hartree–Fock (UHF) approximation, 336–337, 339
- upper-branch polariton, 135
- Urbach rule, 213, 220, 223–224, 333

- vacancy, 359–360
- valence band, 108–109
- valence electron, 284
- valence exciton, 79
- valence mixing, 318
- valley (of energy band), 112
- van der Waals force, 321, 326
- variational method, 155–156
- vector potential, 4, 31
- velocity of light, 3, 133
- vibrational energy quantum, 41–42
- vibrational structures, 48
- vibrationally-allowed transition, 53
- vibronic band, 233
- vibronic exciton, 209, 233
- vibronic line, 209
- vibronic state, 209
- virial theorem, 178
- virtual electron–hole pairs, 166
- virtual excitons, 167
- virtual molecule, 381
- virtual phonon, 153
- virtual process, 247, 251
- V_K-center, 326

- Wannier exciton, 117, 147
- Wannier functions, 108, 283, 391
- water, 371, 383
- wave function, 10, 27
- wave picture, 10
- wave vector, 3
- wavelength, 3
- wave–particle duality, 22
- weak coupling, 217, 219–220, 257, 268
- weak scattering (regime), 141, 201, 219–221
- white excitation, 270
- width function, 193, 197, 221
- Wien's displacement law, 9
- Wolfram's red salt, 353
- wurtzite structure, 132

- X-ray diffraction, 298, 338

- zero-phonon line, 31, 51, 204, 207, 213, 234, 304
- zero-phonon line of exciton, 194
- zero-point energy, 19, 21, 279
- zinc diphosphide, 274
- zinc-blende structure, 324